EINDHOVEN UNIVERSITY OF TECHNOLOGY

Communication Theory (5ETB0)
Combined Slides

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Communication Theory (5ETB0)
Module 1.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Course staff

Prof. Alex Alvarado
Full Professor

Jasper Lagendijk
PhD Student

Felipe Villenas Cortez
PhD Student

Carmen Alvarez Roa
PhD Student

Mohamed Aiham Hemza
MSc. Student

For more info

http://www.sps.tue.nl/ http://www.tue.nl/ictlab/

Before and After 5ETB0

Level (Before)

- Prerequisites: Mathematics II (5EMA0) and Intro telecommunications (5ETA0)
- Expected knowledge in signals and systems: Fourier theory, LTI systems, probabilities, etc.
- 8 Weeks, 5 ECTS. 1 ECTS $\approx$ 30 hours $\rightarrow$ 140 hours course, $\approx$ 18 hours/week

What After 5ETB0?

- 5XSE0 Information Theory (Q3, BSc) (The course is on-hold for the year 2024-2025)
- 5LSF0 Applications of information theory (Q4, MSc)
- 5LSK0 Digital wireless communication exploration lab (Q4, MSc)
- 5LPA0 Wireless Communications (Q2, MSc)
- 5LTB0 Fibre optic communication systems and networks (Q4, MSc)
- 5STA0 Optical fibre communication technology (Q3, MSc)

Coherent Package

Machine Learning and Information Processing for Communications

ICT Lab offers a coherent package covering digital information from telecommunications, information theory, and machine learning perspectives.

Information and communication technologies ID 116557498 ©Funtap P Dreamstime.com

Course Code	Code Name	Schedule
5XSL0	Fundamentals of Machine Learning	Q4 (timeslot D), year 2
5XSE0	Information Theory*	Q3 (timeslot A), year 3
5XTA0	Telecommunication systems	Q4 (timeslot B), year 3

* The course is on hold for the year 2024-2025

https://www.sps.tue.nl/ictlab/education/

Learning Outcomes

After completion of this course, the student should be able to:

- Summarize the advantages and disadvantages of digital communication systems.
- Explain the properties of white Gaussian noise and the differences between wideband, baseband, and passband channels.
- Determine the optimal decision rules and resulting error-probabilities for scalar and vector channels with discrete and real-valued outputs.
- Use the Q-function for evaluating error probabilities.
- Explain the differences between MAP and ML receivers in terms of optimality, complexity and performance.
- Compute the energy of signals and test orthogonality.
- Determine the vector representation of signals by constructing a set of building block waveforms.
- Use Parseval's relationship to solve equivalent time-frequency domain problems.
- Design and implement correlation, matched filter, and direct receivers for the AWGN channel.
- Explain the connection between a matched filter and the resulting signal-to-noise ratio.
- Describe the concept of signal structure and the effects of translation and rotation on the error probability.
- Recognize orthogonal signals, construct optimum Rxs for such signals, and explain the behavior of its error probability when the number of signals increases.
- Evaluate and compare the performance of communication systems in terms of signal-energy, bandwidth (dimensionality), and signal-to-noise ratio.
- Compare the performance of bit-by-bit and block-orthogonal signaling.
- Estimate the number of orthogonal waveforms N (dimensions) that can be fitted into a given bandwidth W.
- Explain the concept of channel capacity and its relation to transmission rate, block length, and error probability.
- Determine and estimate the capacities of band-limited, wide-band, and band-pass-channels.
- Explain the general principles behind the proof of Shannon's channel capacity theorem.
- Explain the importance of the Nyquist criterion and recognize pulses that satisfy this criterion (both in time and frequency domain).
- Describe the role that sinc pulses play in terms of transmission with minimum bandwidth.
- Use cosine and sine functions to transmit information in a passband channel.
- Compute the capacity of the passband channel and compare it to the equivalent baseband channel.
- Compute error probability of QAM constellations using Q-functions.
- Construct and analyze optimal receivers incoherent transmission.
- Explain and implement different blocks that form a standard digital communication system.

Course Elements (1/2)

Parts

- **Lectures (Alex)**: Divided in 12 modules plus 1 invited lecture. Live lectures based on pre-recorded videos.
- **Instructions (TAs)**: 2 per week, each on single module discussed in lecture. Mix of self-study problem solving and problem solutions by TAs. Preparation for **same-day** quiz.
- **Quizzes (TAs)**: 2 per week (except: first week (only on Nov. 13), during Q&A sessions, and the invited lecture)
  2 questions: first, about the lecture/video lecture and second, taken from the exercise bundle. Not mandatory but part of the final grade
- **Q&A Sessions (TAs)**: To discuss modules/quiz/assignment content. If there are no questions, the session will be cancelled
- **Assignments**: 3 assignments, each with a mathematical problems part, and a second part with one MATLAB problem. Part of the final grade
- **Study Guide**: All you need to know about the course
- **Course Reader**: 12 Chapters
- **Written Exam**: Jan. 22 2024 13h30m Mandatory (re-sit on April 12 13h30m)

Course Elements (2/2)

Exercise Bundle

- Contains exercises grouped per chapter
- Answers **and solutions** to all exercises are provided
- Student-driven. Exercises may be explained by TAs if multiple students have the same question.

A Typical Day

- Lecture
- Instruction
- Quiz + Q&A Quiz
- Instruction/Lecture

Course Planning

Course Planning						
Date	Module	Instruction	Quiz	Lecture	Assignment	
					Release	Submission
11 Nov.	M1	-	No	Live		
13 Nov.	M2	M2	Yes (M2)	Half-half		
18 Nov.	M3	M3	Yes (M3)	Half-half	A1	
20 Nov.	M4	M4	Yes (M4)	Half-half		
25 Nov.	M5	M5	Yes (M5)	Half-half		
27 Nov.	M6	M6	Yes (M6)	Half-half		
2 Dec.	Q&A Session	M1-6 + A1	No	No	A2	
3 Dec.	-	-				A1
4 Dec.	M7	M7	Yes (M7)	Half-half		
9 Dec.	M8	M8	Yes (M8)	Half-half		
11 Dec.	M9	M9	Yes (M9)	Half-half		
16 Dec.	M10	M10	Yes (M10)	Half-half		
18 Dec.	Q&A Session	M1-9 + A2	No	No	A3	
19 Dec.	-	-				A2
6 Jan.	M11	M11	Yes (M11)	Half-half		
8 Jan.	M12	M12	Yes (M12)	Half-half		
13 Jan.	Invited Lecture	-	No	Live		
15 Jan.	Q&A Session*	M1-12 + A3	No	No		
16 Jan.	-	-				A3
22 Jan.	**Final Examination**					

* The Q&A Session on the 15 Jan. will only be with the TAs.

Extra Info

Do you need a quick answer?

Take a look at www.wikipedia.org, www.google.com, or ask ChatGPT

Do you need a slow (but better) answer?

Take a look at the reader, additional reading material, or ask a question!

How to Contact Us?

- During Q&A Sessions and Instructions
- Request a meeting via email with any of the course staff
- Visit us in Floor 7 of Flux building

Details

Lectures

- The video lectures cover essentials only. Watch the pre-recorded lectures before coming to the lecture!
- All details in the course reader
- A combination of slides, examples, and MATLAB demos will be used

Instructions

- Solutions can be found in the Exercise Bundle
- Instructions consist of a mix of
  - Guided self-study
  - Solution of selected problems
  - Q&A
- Try to read and solve the problems before the exercise session
- Complete (at home) potentially unfinished self-study exercises
- The level of the problems is often, but not always, similar to the level of the exam's problems

Quizzes and Assignments

Details

- **Quizzes** will consist of simple questions to evaluate your knowledge on the course
- Questions in **quizzes** will be mostly theoretical, but small calculations can be present
- **Quizzes** will be paper-based
- The three lowest score quizzes will not be considered in the quiz grade calculation
- **Quizzes** will be placed after the instructions and will be 15 minute long
- **Assignments** are more elaborated and can consist of programming tasks or more difficult questions involving calculations
- You will have more time (days) to solve **assignments**
- **Assignments** will require MATLAB

Written Exam

Details

- **Understanding** communication theory will be rewarded
- Emphasis is not on memorizing facts or solving standard problems
- Aim for understanding **during the course**. The earlier the preparation starts, the better.
- The **solution** is more important than the **answer**:
  - A good solution with a minor error usually gives close to full points, even if the answer is incorrect
  - An answer without a clear motivation usually gives 0 points, even if it is correct
- **Instructions** and **Q&A Sessions** will help you improve your understanding. Use them!
- Equations (max one-page) will be attached to the exam. File in Canvas already.

Grading and Rules

Grading

- Individual grades
- Final Grade = Exam 70% + Quizzes 10% + Assignments 20%
- Exam grade must be at least 5.0

Remember...

- **Questions** are always very **welcome**
- **Late arrivals** are **not welcome**
- **Interactions** are encouraged during instructions and Q&A sessions

Cheating

- Group discussions are encouraged
- Plagiarism software for code
- Plagiarism will be reported

Summary Module 1.1

Take Home Messages

- Course Overview
- Online Course:
  - Pre-recorded Lectures
  - Q&A Sessions
  - Instructions
- Final Grade:
  - Exam
  - Quizzes
  - Assignments

EINDHOVEN UNIVERSITY OF TECHNOLOGY

ICT LAB

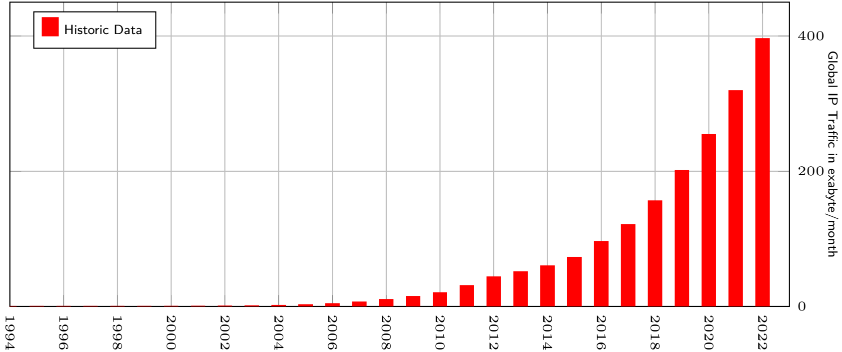
Communication Theory (5ETB0)
Module 1.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Course Motivation (1/4)



Global IP traffic per year according to the Cisco virtual networking index in exabyte/month (1 exabyte$=10^{18}$ byte).

- Wireless (WiFi, 4G, DVB), wired (coax, twisted pair), fiber optics, etc.
- Study underlying theory and principles behind the global telecomms infrastructure

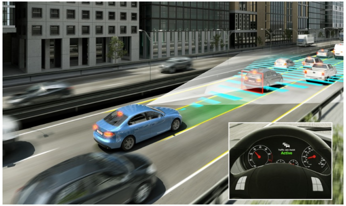
Course Motivation (2/4)

iCAVE Project: Embedding Digital Communication in Car Radars

Context:
Vehicle-to-Vehicle (V2V) communication and automotive radars are expected to play a strategic role in improving driving safety in future intelligent transportation system (ITS).

The Project:
How can we modify the waveforms use in radar so that they can also be used to transmit information between cars?

Required Knowledge:
Digital communications, modulation, detection theory, etc.



Course Motivation (3/4)

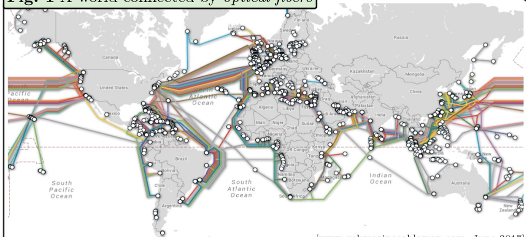
FUN-NOTCH: Fundamentals of the Nonlinear Optical Channel

Context:
Fibre optics are critical infrastructure for society because they carry nearly all the global Internet traffic.

The Project:
Find the maximum amount of information that optical fibres can carry in the highly nonlinear regime and design transceivers well-suited for this regime.

Required Knowledge:
Digital communications, information theory, error correcting codes, etc.

Fig. 1 A world connected by *optical fibers*



[www.submarinecablemap.com, June 2017]

Course Motivation (4/4)

Historic Perspective

Discipline				
Telegraphy and Telephony	Wireless Communications	Electronics	Modulation Methods	Fundamentals
Volta, 1800 Oersted, 1819 Henry, 1827 Morse, 1838 Graham Bell, 1875 Pupin, 1900	Faraday, 1831 Maxwell, 1873 Hertz, 1886 Branly, 1890 Popov, 1896 Marconi, 1901 Pierce, 1955	Braun, 1874 Fleming, 1904 DeForest,1906 Bardeebm Brattain, Shockley, 1948 Kilby, Noyce, 1958	Campbell, 1909 Armstrong, 1915-1918 Bijl, Hartley, Heising, 1915 Carson, 1915	Nyquist, 1924 Hartley, 1928 Reeves, 1938 Dudley, 1939 Wiener, 1942 **North, 1943** Kotelnikov, 1947 **Shannon, 1948**

- North invented the matched filter
- Shannon is known as “the father of information theory”
- More details in Chapter 1 of the reader
- A (relatively speaking) new field

Summary Module 1.2

Take Home Messages

- Theory behind digital data transmission
- Exciting field with many interesting applications

Communication Theory (5ETB0)
Module 2.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 2.1

Presentation Outline

Signals and Systems Review: Fourier Transform (Appendix A)

The Fourier Transform

The Fourier Transform pair is defined as

$$\mathcal{F}\{x(t)\} = X(f) \triangleq \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-\jmath 2\pi f t}\, dt \iff x(t) = \int_{-\infty}^{\infty} X(f)\mathrm{e}^{\jmath 2\pi f t}\, df,$$

or alternatively,

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-\jmath\omega t}\, dt \iff x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\mathrm{e}^{\jmath\omega t}\, d\omega.$$

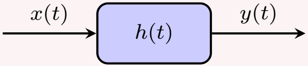
Fourier Transform Properties

- Linearity, i.e., $\mathcal{F}\{ax_1(t) + bx_2(t)\} = aX_1(f) + bX_2(f)$
- Transform of a convolution: $\mathcal{F}\{x_1(t) * x_2(t)\} = X_1(f) \cdot X_2(f)$
- If $x(t) \in \mathbb{R}$, its Fourier transform satisfies $X(f) = X^*(-f)$
- If $x(t) \in \mathbb{R}$ and even (i.e., symmetric respect to zero: $x(-t) = x(t)$) its Fourier transform is real ($X(f) \in \mathbb{R}$) and even ($X(f) = X(-f)$)

Signals and Systems Review: LTI System (Appendix C)

An LTI system

- The impulse response of the LTI system is given by $h(t)$
- In the time domain, $y(t) = x(t) * h(t)$
- In the frequency domain, $Y(f) = X(f)H(f)$



An LTI system. The output $y(t)$ is the convolution between the input $x(t)$ and the impulse response $h(t)$.

Signals and Systems Review: Parseval Relation (Appendix A)

Parseval's Theorem

Parseval's Relation states that for two real signals $g_1(t)$ and $g_2(t)$

$$\int_{-\infty}^{\infty} g_1(t)g_2(t)\,dt = \int_{-\infty}^{\infty} G_1(f)G_2^*(f)\,df,$$

which for the particular case of $g_1(t) = g_2(t) = g(t)$ translates into

$$\int_{-\infty}^{\infty} g^2(t)\,dt = \int_{-\infty}^{\infty} |G(f)|^2\,df,$$

where

$$E_g \triangleq \int_{-\infty}^{\infty} g^2(t)\,dt$$

is the energy of the signal.

Module 2.1

Presentation Outline

Mathematical Notation (1/3)

- *Sets* are denoted by calligraphic letters: $\mathcal{M}$

- The *cardinality* of a set is denoted by $|\mathcal{M}|$

- A *definition* is denoted by $\triangleq$

Mathematical Notation (2/3)

- A set definition is therefore: $\mathcal{M} \triangleq \{1, 2, \cdots, |\mathcal{M}|\}$
- Exceptions include the sets of real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$
- Cartesian product of sets: $\mathcal{M}^2 = \mathcal{M} \times \mathcal{M}$
- Notation $\mathcal{M} \subset \mathbb{R}$ means $\mathcal{M}$ is a subset of $\mathbb{R}$

Mathematical Notation (3/3)

- *Estimated* variables are denoted using a hat: $\hat{m}$
- *Scalars* are denoted by small letters: x
- *Vectors* are denoted using underlined letters: $\underline{x}$
- The function $\min_{m\in\mathcal{M}}\{\cdot\}$ is not the same as $\mathrm{argmin}_{m\in\mathcal{M}}\{\cdot\}$

Module 2.1

Presentation Outline

Discrete Random Variables (1/4)

- Random Variables (RVs) are denoted by capital letters and their realizations by small letters: X is not the same as x

- $\Pr\{X = x\} \geq 0$ denotes the probability that the r.v. X takes the value x

- The support of the r.v. X will be denoted by a set: $\mathcal{X} = \{x_1, x_2, \ldots, x_{|\mathcal{X}|}\}$

Discrete Random Variables (2/4)

- The probability mass function (PMF) of a random variable is denoted by $\Pr\{X = x\}$ for all $x \in \mathcal{X}$.

- PMFs satisfy

$$\sum_{x \in \mathcal{X}} \Pr\{X = x\} = 1$$

Discrete Random Variables (3/4)

- Joint PMFs are denoted by $\Pr\{X = x, Y = y\}$
- Conditional PMFs are denoted by $\Pr\{Y = y|X = x\}$.
- Conditional PMFs satisfy

$$\begin{aligned} \Pr\{Y = y, X = x\} &= \Pr\{Y = y|X = x\} \cdot \Pr\{X = x\} \\ &= \Pr\{X = x|Y = y\} \cdot \Pr\{Y = y\} \end{aligned}$$

Discrete Random Variables (4/4)

- Bayes' Rule:

$$\Pr\{Y = y|X = x\} = \frac{\Pr\{X = x|Y = y\} \cdot \Pr\{Y = y\}}{\Pr\{X = x\}}$$

- Law of total probability:

$$\begin{aligned}\Pr\{Y = y\} &= \sum_{x\in\mathcal{X}} \Pr\{X = x, Y = y\} \\ &= \sum_{x\in\mathcal{X}} \Pr\{Y = y|X = x\} \cdot \Pr\{X = x\}\end{aligned}$$

- Independence

$$\Pr\{Y = y|X = x\} = \Pr\{Y = y\}$$

Module 2.1

Presentation Outline

Continuous Random Variables (1/2)

- Just like for discrete RVs, continuous RVs are denoted by capital letters and their realizations by small letters: R and r

- The support of the random variable is denoted by a calligraphic letter $\mathcal{R}$ ($\mathbb{R}$ for real numbers)

- The probability density function (PDF) of a random variable is $p_R(r)$ for all $r \in \mathcal{R}$

- $p_R(r)dr$ denotes the probability that the r.v. R takes a value between r and $r + dr$

Continuous Random Variables (2/2)

- PDFs satisfy

$$\int_{r\in\mathcal{R}} p_R(r)dr = 1$$

- Joint PDFs are denoted by $p_{R,S}(r,s)$ and also satisfy

$$\int_{r\in\mathcal{R}}\int_{s\in\mathcal{S}} p_{R,S}(r,s)dsdr = 1$$

- Conditional PDFs are $p_R(r|X=x)$. Law of total probability is

$$p_R(r) = \sum_{x\in\mathcal{X}} \Pr\{X=x\}\cdot p_R(r|X=x)$$

Summary Module 2.1

Take Home Messages

- Often used in the course:
  - Fourier Transforms
  - Convolutions
  - LTI Systems
- Notation Convention
- Discrete and Continuous Random Variables

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

ICT LAB

Communication Theory (5ETB0)
Module 2.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 2.2

Presentation Outline

The General Communication Problem

Elements in a communications system

Source $\xrightarrow{u(t)}$ Transmitter $\xrightarrow{s(t)}$ Channel $\xrightarrow{r(t)}$ Receiver $\xrightarrow{v(t)}$ Destination

- **Source**: Emits a real-valued waveform which contains the information to be transmitted (mic., sensor, etc.)
- **Transmitter**: Converts the source waveform into the waveform that will be sent through the channel (distance)
- **Channel**: Accepts a waveform and it gives as output another waveform (continuous- or discrete-input)
- **Receiver**: Tries to reconstruct $u(t)$ based on $r(t)$ (guessing)

Module 2.2

Presentation Outline

White Gaussian Noise

White Noise Waveform

A white noise waveform $n_w(t)$ is generated by the random process $N_w(t)$, which is (i) zero-mean, (ii) stationary, (iii) white, and (iv) Gaussian.

(i) Zero mean: $E[N_w(t)] = 0$

(ii) Stationary:

$$R_{N_w}(t,s) \triangleq E[N_w(t)N_w(s)] = \frac{N_0}{2}\delta(t-s),$$

(iii) White:

$$S_{N_w}(f) = \frac{N_0}{2}\left[\frac{W}{Hz}\right], \text{ for } -\infty < f < \infty,$$

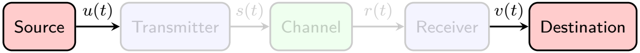
where $S_{N_w}(f)$ is the PSD, i.e, the FT of the autocorr. function.

(iv) Gaussian: The random vector of samples

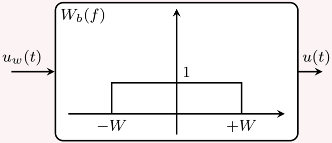
$$\underline{n} = (N_w(t_1), N_w(t_2), \dots, N_w(t_N))$$

is jointly Gaussian for any finite set of sampling instants $\{t_1, t_2, \dots, t_N\}$

A Waveform Source



Baseband Gaussian Source

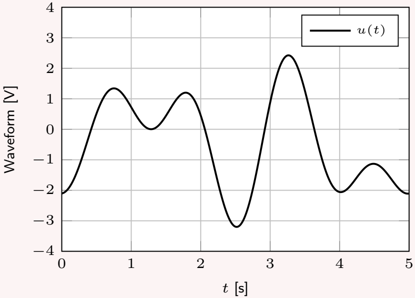


$$u(t) = u_w(t) * w_b(t)$$

Limitations:

- Not the best model. Analog sources typically not white, not baseband, not Gaussian
- Worst source we can have (in terms of compression)

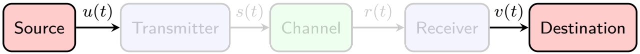
Example source waveform $u(t)$



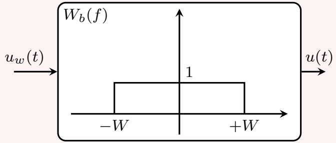
Bandwidth is $W = 1$ Hz.

How different would it be if $W = 10$ Hz?

A Waveform Source



Baseband Gaussian Source

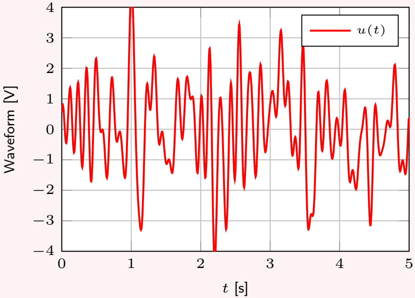


$$u(t) = u_w(t) * w_b(t)$$

Limitations:

- Not the best model. Analog sources typically not white, not baseband, not Gaussian
- Worst source we can have (in terms of compression)

Example source waveform $u(t)$

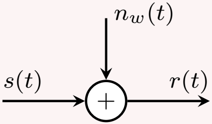


Bandwidth is $W = 1$ Hz.

How different would it be if $W = 10$ Hz?

Three Waveform Channels (1/2)

(1) Wideband AWGN Channel

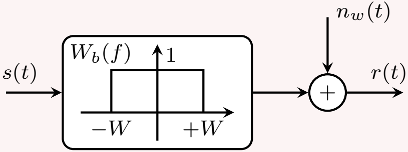


Wideband AWGN channel

$$r(t) = s(t) + n_w(t)$$

Why is this not realistic?

(2) Baseband AWGN Channel



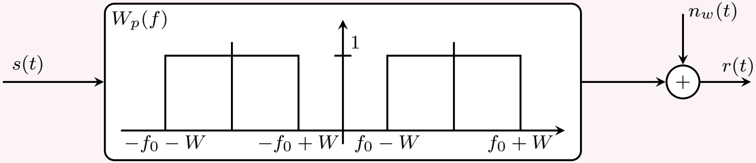
Baseband AWGN channel

$$r(t) = s(t) * w_b(t) + n_w(t)$$

Good model for copper cables

Three Waveform Channels (2/2)

(3) Passband AWGN Channel



Passband AWGN channel

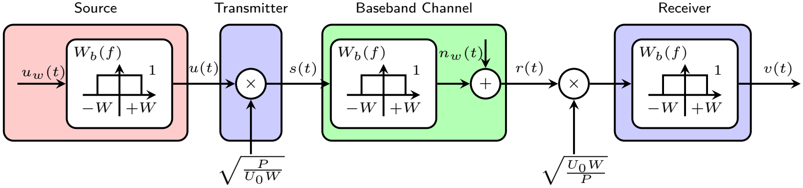
$$r(t) = s(t) * w_p(t) + n_w(t),$$

Passband transmission with frequency determined by the channel

Module 2.2

Presentation Outline

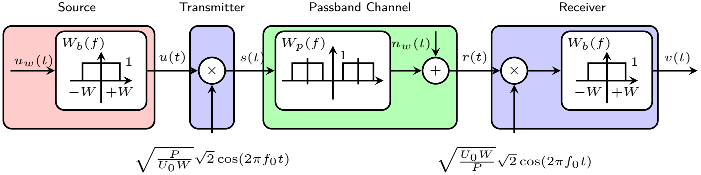
Baseband Transmission



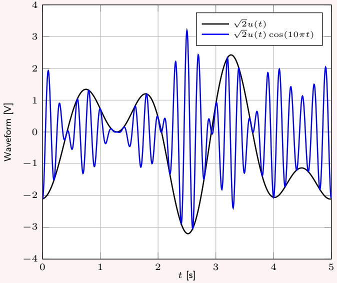
Baseband Transmission

- **Definition:** *Instantaneous power is* $s^2(t)$. *Average power is* $E[S^2(t)]$. *Average transmit power* P *is limited, and thus,* $E[S^2(t)] \leq P$.
- **Power distortion**: $d(t) \triangleq v(t) - u(t) = \sqrt{U_0 W/P} \cdot n_w(t) * w_b(t)$
- **Variance distortion**: $E[D^2(t)] = \frac{U_0 W}{P} \frac{N_0}{2} 2W = U_0 W \frac{N_0 W}{P}$
- **Mean squared error distortion**: $D = E[D^2(t)]$
- **Source signal-to-distortion ratio**: SDR $\triangleq \frac{E[U^2(t)]}{E[D^2(t)]} = \frac{P}{N_0 W} = \text{SNR}_b$, where SNR_b is the baseband channel **signal-to-noise ratio**.

Passband Transmission (DSB-SC)



Input waveform



SDR for DSB-SC

- The source SDR (signal to distortion ratio)

$$\frac{E[U^2(t)]}{E[D^2(t)]} = \frac{P}{N_0 W}$$

  which is equal to SNR_b.

- *The SDR is the same we found for baseband Tx!*

Pros and Cons of Analog Transmission

Pros:

- Easy to implemente using analog electronics
- Graceful degradation
- Easy multiplexing using frequency-division multiplexing (FDMA)

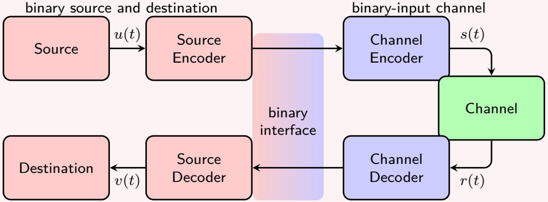
Cons:

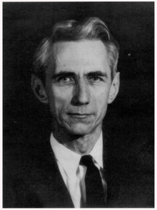
- Not flexible with channel BW requirement
- SNR decrease in each hop (relay)
- Encryption and compression very difficult

Digital Communication

Two key ideas from Shannon (1948)

- All sources can be represented by binary sequences
- Source and channel processing can be separated





Claude Shannon—Father of the Information Age

Two videos in Canvas:

https://youtu.be/z2Whj_nL-x8

https://youtu.be/E3OldEtfBrE

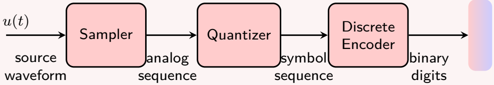
Digital Communication

The Pros of Digital Communication

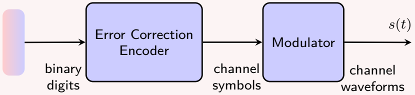
- Cheap, reliable and miniaturized digital hardware
- Simple quality control (error rates, detection, correction)
- Simplified system development thanks to binary interface
- No performance loss by source and channel separation
- Simplified networking thanks to binary interface
- Efficient utilization of resources (source coding)

Source and Channel Processing

Dividing Source and Channel Processing



Dividing source processing at the transmit side into sampling, quantizing, and discrete encoding



Dividing channel processing at the transmit side into error correction encoding, and modulation

Source processing and error correction not treated in this course, but in

- 5XSE0 Information Theory (Q3, BSc)
- 5LSF0 Applications of information theory (Q4, MSc)
- In the invited lecture

Summary Module 2.2

Take Home Messages

- Communication problem: a guessing problem
- Analog sources, white Gaussian noise
- Baseband Gaussian source
- Three channels: Wideband, Baseband and Passband AWGN
- Digital vs. Analog? Digital!

Communication Theory (5ETB0)
Module 3.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

Module 3.1

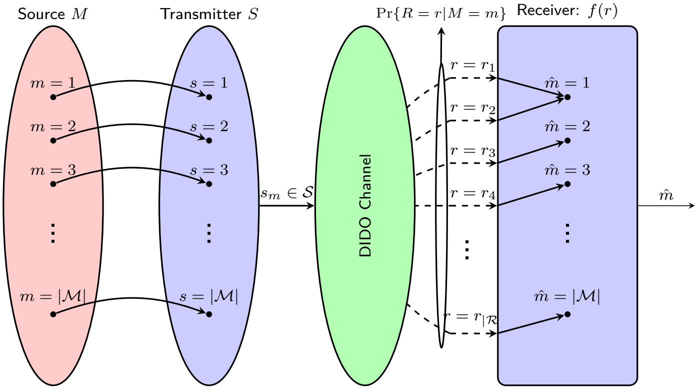
Presentation Outline

Definitions (1/2)

Definitions

- Source: Produces a *message* $m \in \mathcal{M} \triangleq \{1, 2, \ldots, |\mathcal{M}|\}$ with probability $\Pr\{M = m\}$ for $m \in \mathcal{M}$. The r.v. is M
- Transmitter: Sends a *signal* $s_m \in \mathcal{S}$ if message m is to be transmitted. The r.v. is S
- Channel: Produces output $r \in \mathcal{R}$ (r.v. is R) with conditional probability $\Pr\{R = r|S = s\}$
- Receiver: Forms an *estimate* $\hat{m}$ by observing the received channel output $r \in \mathcal{R}$ using a mapping $\hat{m} = f(r) \in \mathcal{M}$. The r.v. is $\hat{M}$

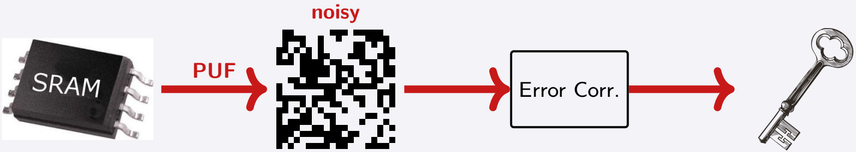
Definitions (2/2)



Motivation for DIDO Channels: SRAM-PUF

Average Error Probability

- Use power-on value of SRAM to generate cryptographic keys
- Error-Correcting Codes required to reliably reconstruct the key from noisy binary values
- RESCURE project: improve security, reliability, performance



Rescure
eurostars™
Co-funded by EUREKA member countries
and the European Union Horizon 2020
Framework Programme

Module 3.1

Presentation Outline

Error Probability Definitions

The Detection Problem

- For a given channel, find the best **decision rule** $f(r)$
- Best in what sense? Error probability...

Average Error Probability

The **probability of error** is defined as

$$P_{\mathsf{e}} \triangleq \Pr\{\hat{M} \neq M\}. \tag{1}$$

The **probability of correct decision** is defined as

$$P_{\mathsf{c}} \triangleq \Pr\{\hat{M} = M\} = 1 - P_{\mathsf{e}}. \tag{2}$$

Optimum Receiver

A receiver is optimum if it minimizes the error probability P_{e}.

Correct Probability via Joint PMF (1/2)

Average Error Probability

The correct probability can be expressed as

$$\begin{aligned}
P_{\mathsf{c}} &= \Pr\{M = \hat{M}\} \\
&= \Pr\{M = f(R)\} \\
&= \sum_{r \in \mathcal{R}} \Pr\{R = r, M = f(r)\} \\
&= \sum_{r \in \mathcal{R}} \sum_{m \in \mathcal{M}} \Pr\{R = r, m = f(r) | M = m\} \Pr\{M = m\} \\
&= \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}} \Pr\{R = r, m = f(r) | M = m\} \Pr\{M = m\} \\
&= \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}: f(r) = m} \Pr\{R = r | M = m\} \Pr\{M = m\} \\
&= \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}: f(r) = m} \Pr\{M = m, R = r\}
\end{aligned}$$

Correct Probability via Joint PMF (2/2)

Error Probability Computation: A Recipe

- We showed that:

$$P_{\mathsf{c}} = \sum_{m \in \mathcal{M}} \sum_{r \in \mathcal{R}: f(r)=m} \Pr\{M=m, R=r\}$$

- Make a table with $\Pr\{M=m, R=r\}$ for all possible combinations of m and r
- For each $M=m$, find all columns where $f(r)=m$, and sum them up
- Alternatively, for each $R=r$, identify the entry in the table that the detection rule $f(r)$ will choose

Example 3.1 (1/4)

Applying the Recipe

- Tx signals: $s \in \mathcal{S} = \{s_1, s_2\}$ $(|\mathcal{M}| = 2)$. Rx signals: $r \in \mathcal{R} = \{a, b, c\}$
- A-priori probabilities:

m	$\Pr\{M = m\}$
1	0.4
2	0.6

- Conditional probabilities:

m	$\Pr\{R = a\|S = s_m\}$	$\Pr\{R = b\|S = s_m\}$	$\Pr\{R = c\|S = s_m\}$
1	0.5	0.4	0.1
2	0.1	0.3	0.6

- Decision rule: $f(r) = 1$ if $r \in \{a, b\}$ and $f(r) = 2$ if $r = c$
- Joint probabilities:

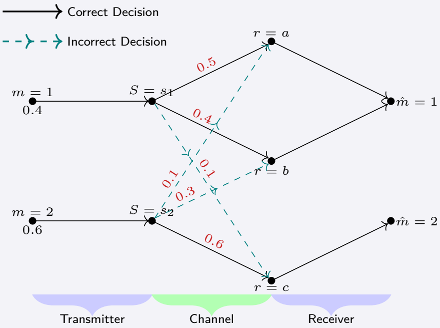
m	$\Pr\{M = m, R = a\}$	$\Pr\{M = m, R = b\}$	$\Pr\{M = m, R = c\}$
1	0.20	0.16	0.04
2	0.06	0.18	0.36

- Correct probability is $P_{\mathrm{c}} = 0.2 + 0.16 + 0.36 = 0.72 \Rightarrow P_{\mathrm{e}} = 0.28$

Example 3.1 (2/4): A different view

A Graphical Interpretation

$$P_e = \Pr\{\hat{M} \neq M\} = \sum_{m \in \mathcal{M}} \Pr\{\hat{M} \neq M | M = m\} \Pr\{M = m\} \quad (3)$$



Error probability is $P_e = 0.4 \cdot 0.1 + 0.6 \cdot (0.1 + 0.3) = 0.28 \Rightarrow P_c = 0.72$

Module 3.1

Presentation Outline

Example 3.1 (3/4)

Maximum Likelihood Detection

- Can we increase P_c by improving $f(r)$?

$$\boxed{P_\mathsf{c} = \sum_{m\in\mathcal{M}} \sum_{r\in\mathcal{R}: f(r)=m} \Pr\{M=m, R=r\}}$$

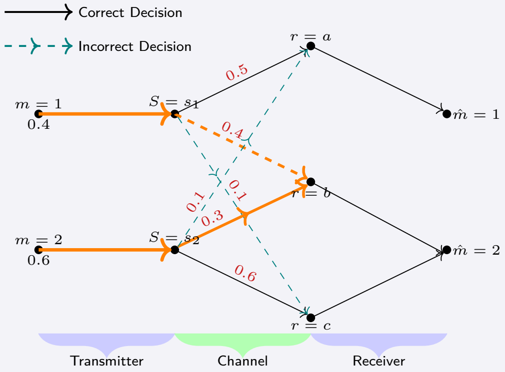
- For each column, the decision rule picks one row
- The example:

m	$\Pr\{M=m,R=a\}$	$\Pr\{M=m,R=b\}$	$\Pr\{M=m,R=c\}$
1	0.20	0.16	0.04
2	0.06	0.18	0.36

- A higher correct probability is $P_\mathsf{c} = 0.2+0.18+0.36 = 0.74 \Rightarrow P_\mathsf{e} = 0.26$

Example 3.1 (4/4): A different view

A Graphical Interpretation



Error probability is $P_{\text{e}} = 0.4 \cdot (0.1 + 0.4) + 0.6 \cdot 0.1 = 0.26 \Rightarrow P_{\text{c}} = 0.74$

Summary Module 3.1

Take Home Messages

- DIDO Channels and problem definition
- Error probability definition and calculations
- Detection can be improved

Communication Theory (5ETB0)
Module 3.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 3.2

Presentation Outline

MAP Detection (1/2)

Motivation MAP Decision Rule

- Recall that

$$P_{\mathsf{c}} = \sum_{r \in \mathcal{R}} \Pr\{R = r, M = f(r)\} \tag{4}$$

- Interpretation: For each column $(R = r)$, decision rule picks a row $(M = m)$.
- This interpretation leads to the upper bound

$$P_{\mathsf{c}} \leq \sum_{r \in \mathcal{R}} \max_{m \in \mathcal{M}} \Pr\{R = r, M = m\} \tag{5}$$

- Upper bound is achieved by $f(r)$ that picks the row that maximizes the joint probability

MAP Detection (2/3)

A Different Interpretation

The optimum receiver is a maximization over $f : \mathcal{R} \to \mathcal{M}$, i.e.,

$$\max_f \{P_c\} = \max_f \sum_{r \in \mathcal{R}} \Pr\{R = r, M = f(r)\} \tag{6}$$

$$= \sum_{r \in \mathcal{R}} \max_{f(r)} \Pr\{R = r, M = f(r)\} \tag{7}$$

$$= \sum_{r \in \mathcal{R}} \max_{f(r)} \Big\{ \underbrace{\Pr\{R = r, 1 = f(r)\}, \ldots, \Pr\{R = r, |\mathcal{M}| = f(r)\}}_{\begin{cases} \Pr\{R = r, M = m\}, & \text{if } f(r) = m \\ 0, & \text{if } f(r) \neq m \end{cases}} \Big\} \tag{8}$$

Thus,

$$\max_f \{P_c\} = \sum_{r \in \mathcal{R}} \max_{m \in \mathcal{M}} \Pr\{R = r, M = m\} \tag{9}$$

$$= \sum_{r \in \mathcal{R}} \max_{m \in \mathcal{M}} \Pr\{M = m | R = r\} \xcancel{\Pr\{R = r\}} \tag{10}$$

MAP Detection (3/3)

Decision Variables

For a communication system using a DIDO channel, the joint PMFs

$$\begin{aligned} \Pr\{M=m, R=r\} &= \Pr\{M=m\}\Pr\{R=r|M=m\} && (11) \\ &= \Pr\{M=m\}\Pr\{R=r|S=s_m\} && (12) \end{aligned}$$

are called the **decision variables**. An optimum receiver uses these variables.

MAP Decision Rule

The maximum a-posteriori probability (MAP) decision rule is

$$\hat{m}^{\mathsf{MAP}}(r) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{M=m|R=r\} \qquad (13)$$

and has two important properties:

- $\Rightarrow$ Maximizes $P_{\mathsf{c}} \Rightarrow$ Minimizes $P_{\mathsf{e}} \Rightarrow$ optimum receiver!
- Produces the largest decision variable for each r (Bayes' rule)

Example 3.1 Revisited

MAP for Example 3.1

- A-posteriori probabilities:

m	$\Pr\{M=m\mid R=a\}$	$\Pr\{M=m\mid R=b\}$	$\Pr\{M=m\mid R=c\}$
1	20/26	16/34	4/40
2	6/26	18/34	36/40

  obtained from $\Pr\{R=a\}=0.26, \Pr\{R=b\}=0.34$, and $\Pr\{R=c\}=0.4$
- Correct probability is $P_\mathsf{c}=0.2+0.18+0.36=0.74 \Rightarrow P_\mathsf{e}=0.26$
- MAP decision rule coincides with the one that maximizes P_c

Intuition behind MAP decision rule

Maximize the probability that, for a given $R=r$, the chosen message is equal to the transmitted message

$$\hat{m}^{\mathsf{MAP}}(r) \stackrel{\Delta}{=} \operatorname*{argmax}_{m\in\mathcal{M}} \Pr\{M=m\mid R=r\}$$

Module 3.2

Presentation Outline

Detection with Equally Likely Messages

Probabilities

- A-priori probabilities: $\Pr\{M = m\}$
- A-posteriori probabilities: $\Pr\{M = m | R = r\}$

Uniform a-priori Probabilities

- All messages are equally likely (uniform probability)

$$\Pr\{M = m\} = \frac{1}{|\mathcal{M}|} \text{ for all } m \in \mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}, \tag{14}$$

- Decision variables are

$$\Pr\{M = m, R = r\} = \Pr\{R = r | S = s_m\} \Pr\{S = s_m\} \tag{15}$$

$$= \frac{1}{|\mathcal{M}|} \Pr\{R = r | S = s_m\} \tag{16}$$

ML Detection

ML decision rule

The maximum likelihood (ML) decision rule is

$$\hat{m}^{\mathsf{ML}}(r) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{R = r | M = m\}$$

A few words on ML

- Name comes from definition
- For equally likely messages:
  - Largest decision variable for each r
  - $\Rightarrow$ Maximizes P_{c} $\Rightarrow$ Minimizes P_{e} $\Rightarrow$ optimum receiver!
- For nonequally likely messages:
  - Can be used (it is simple to implement)
  - Suboptimal

Example 3.1 Re-revisited (1/2)

ML for Example 3.1

- Transition probabilities:

m	$\Pr\{R=a\|S=s_m\}$	$\Pr\{R=b\|S=s_m\}$	$\Pr\{R=c\|S=s_m\}$
1	0.5	0.4	0.1
2	0.1	0.3	0.6

- The maximum-likelihood decision rule is then:

r	a	b	c
$f(r)$	1	1	2

  Correct probability is 0.72, lower than MAP (0.74).

- If a-priori probabilities are $\Pr\{M=m\}=1/2$:

$$P_{\mathsf{c}}^{\mathsf{ML}} = \Pr\{\hat{M}^{\mathsf{ML}} = M\} \tag{17}$$

$$= \sum_{m\in\mathcal{M}} \Pr\{\hat{M}^{\mathsf{ML}} = M|M=m\}\Pr\{M=m\} \tag{18}$$

$$= \frac{1}{2}(0.5+0.4) + \frac{1}{2}0.6 = 0.75. \tag{19}$$

Example 3.1 Re-revisited (2/2)

ML for Example 3.1

- Transition probabilities:

m	$\Pr\{R = a \mid S = s_m\}$	$\Pr\{R = b \mid S = s_m\}$	$\Pr\{R = c \mid S = s_m\}$
1	0.5	0.4	0.1
2	0.1	0.3	0.6

- The maximum-likelihood decision rule is then:

r	a	b	c
$f(r)$	1	1	2

- If a-priori probabilities are $\Pr\{M = m\} = 1/2$:

$$P_{\mathrm{c}}^{\mathrm{ML}} = \frac{1}{2}(0.5 + 0.4) + \frac{1}{2}0.6 = 0.75. \tag{20}$$

Three Questions

- Q1: What would the MAP detector give?
- Q2: If the a-priori probabilities are now 0.4 and 0.6, $P_{\mathrm{c}}^{\mathrm{ML}} = 0.72$ and $P_{\mathrm{c}}^{\mathrm{MAP}} = 0.74$. Does this make sense?
- Q3: Why do we care about ML if it is suboptimal?

Module 3.2

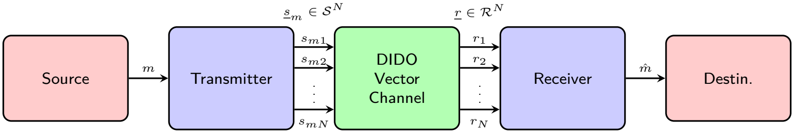
Presentation Outline

MAP and ML for DIDO Channels

Definitions

- Source: Produces a *message* $m \in \mathcal{M} \triangleq \{1, 2, \dots, |\mathcal{M}|\}$ with probability $\Pr\{M = m\}$ for $m \in \mathcal{M}$. The r.v. is M
- Transmitter: Sends a *signal* $s_m \in \mathcal{S}$ if message m is to be transmitted. The r.v. is S
- Channel: Produces output $r \in \mathcal{R}$ (r.v. is R) with conditional probability $\Pr\{R = r | S = s\}$
- Receiver: Forms an *estimate* $\hat{m}$ by observing the received channel output $r \in \mathcal{R}$ using a mapping $\hat{m} = f(r) \in \mathcal{M}$. The r.v. is $\hat{M}$

MAP and ML Detection for Vectorial Channels



Definitions

- Transmitter: Sends a *signal* $\underline{s}_m \in \mathcal{S}^N$ if message m is to be transmitted. The random vector is $\underline{S}$
- Vector Channel: Produces output $\underline{r} \in \mathcal{R}^N$ (random vector is $\underline{R}$) with conditional probability $\Pr\{\underline{R} = \underline{r}|\underline{S} = \underline{s}\}$
- Receiver: Forms an *estimate* $\hat{m}$ by observing the received channel output $\underline{r} \in \mathcal{R}^N$ using a mapping $\hat{m} = f(\underline{r}) \in \mathcal{M}$. The r.v. is $\hat{M}$

MAP and ML Detection: Summary

MAP Detection

Decision	MAP
Variable	$\Pr\{M = m\} \Pr\{\underline{R} = \underline{r} \vert \underline{S} = \underline{s}_m\}$
Rule	$\operatorname{argmax}_{m \in \mathcal{M}} \Pr\{M = m \vert \underline{R} = \underline{r}\}$

ML Detection

Decision	ML
Variable	$\frac{1}{\vert\mathcal{M}\vert} \Pr\{\underline{R} = \underline{r} \vert \underline{S} = \underline{s}_m\}$
Rule	$\operatorname{argmax}_{m \in \mathcal{M}} \Pr\{\underline{R} = \underline{r} \vert M = m\}$

Summary Module 3.2

Take Home Messages

- MAP is the optimal receiver
- ML is sometimes optimal and in general simpler to implement
- Scalar analysis can be generalized to vectorial channels

EINDHOVEN UNIVERSITY OF TECHNOLOGY

ICT LAB

Communication Theory (5ETB0)
Module 4.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 4.1

Presentation Outline

Definitions

Definitions

- Source: Produces a *message* $m \in \mathcal{M} \triangleq \{1, 2, \ldots, |\mathcal{M}|\}$ with probability $\Pr\{M = m\}$ for $m \in \mathcal{M}$. The r.v. is M
- Transmitter: Sends a *signal* $s_m \in \mathcal{S} \subset \mathbb{R}$ if message m is to be transmitted. The r.v. is S
- DICO Channel: Produces output $r \in (-\infty, \infty) = \mathbb{R}$ (r.v. is R) with probability density function $p_R(r|S = s_m) = p_R(r|M = m)$
- Receiver: Forms an *estimate* $\hat{m}$ by observing the received channel output $r \in \mathbb{R}$ using a mapping $\hat{m} = f(r) \in \mathcal{M}$. The r.v. is $\hat{M}$

Decision Variables, MAP and ML

Decision Variables for DICO Channels

The **decision variables** are
$\Pr\{M = m, R = r\} = \Pr\{M = m\} p_R(r|S = s_m) = \Pr\{M = m\} p_R(r|M = m)$

MAP decision rule

The maximum a-posteriori probability (MAP) decision rule is

$$\hat{m}^{\mathsf{MAP}}(r) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{M = m | R = r\} \tag{21}$$

$$= \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{M = m\} p_R(r|M = m) \tag{22}$$

ML decision rule

The maximum likelihood (ML) decision rule is

$$\hat{m}^{\mathsf{ML}}(r) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} p_R(r|M = m) \tag{23}$$

Module 4.1

Presentation Outline

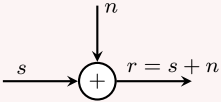
The AGN Channel

Scalar AGN channel

The **AGN channel** adds Gaussian noise N to the input signal S.

This Gaussian noise N has variance σ^2 and mean 0.
Its PDF is

$$p_N(n) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{n^2}{2\sigma^2}\right)$$



The noise variable N is assumed to be independent of the signal S.

Two Questions

$$\text{Q1: } \int_{-\infty}^{\infty} p_N(n)\, dn = ? \quad \text{and} \quad \text{Q2: } p_N(n|S = s_m) \stackrel{?}{=} p_N(n) \tag{24}$$

The AGN Channel: A Matlab Example

Conditional AGN PDF

The conditional PDF of $R = r$ when the signal is $S = s_m$ is a Gaussian PDF, i.e.,

$$p_R(r|S = s_m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - s_m)^2}{2\sigma^2}\right) \tag{25}$$

where σ^2 is the variance of the AGN.

Error Probability

For the Matlab example, if a threshold ad $r^* = 0$ is used, the error probability can be obtained by solving the following integral

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - 1)^2}{2\sigma^2}\right) dr \tag{26}$$

This type of integral is very popular.

Module 4.1

Presentation Outline

MAP Rule for the bi-AGN Channel (1/2)

Two Messages: Binary-input AGN (bi-AGN) Channel

- Assume that $|\mathcal{M}| = 2$: two messages. M can be either 1 or 2.
- The conditional PDF of $R = r$ when the signal is $S = s_m$ is

$$p_R(r|S = s_m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - s_m)^2}{2\sigma^2}\right), \text{ for } m = 1, 2$$

MAP Rule bi-AGN Channel

MAP receiver for $\hat{m} = f(r) = 1$ if

$$\Pr\{M = 1\} p_R(r|S = 1) \geq \Pr\{M = 2\} p_R(r|S = 2) \tag{27}$$

and $\hat{m} = 2$ otherwise. This means $\hat{m} = f(r) = 1$ if

$$\Pr\{M = 1\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - s_1)^2}{2\sigma^2}\right) \geq \Pr\{M = 2\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - s_2)^2}{2\sigma^2}\right)$$

and $\hat{m} = 2$ otherwise.

MAP Rule for the bi-AGN Channel (2/2)

MAP Threshold

One can show (complete derivation in course reader) that $\geq$ becomes $=$ in the MAP ir $r = r^*$, where

$$r^* \triangleq \frac{\sigma^2}{s_1 - s_2} \ln \frac{\Pr\{M = 2\}}{\Pr\{M = 1\}} + \frac{s_1 + s_2}{2}$$

Optimum receiver for the bi-AGN channel

A receiver that decides $\hat{m} = 1$ if

$$r \geq r^*$$

and $\hat{m} = 2$ otherwise, is optimum.

What is needed by MAP?

MAP threshold splits the real line (two intervals) and it depends on noise variance, a-priori probabilities, and transmitted symbols.

ML Rule for the bi-AGN Channel

Optimum Threshold for Equiprobable Messages: ML Rule

When the a-priori probabilities $\Pr\{M=1\}$ and $\Pr\{M=2\}$ are equal, the optimum threshold is

$$r^* = \frac{s_1+s_2}{2}.$$

This corresponds to the ML receiver, which chooses $\hat{m}=f(r)=1$ if

$$p_R(r|S=1) \geq p_R(r|S=2)$$

and $\hat{m}=2$ otherwise. This means $\hat{m}=f(r)=1$ if

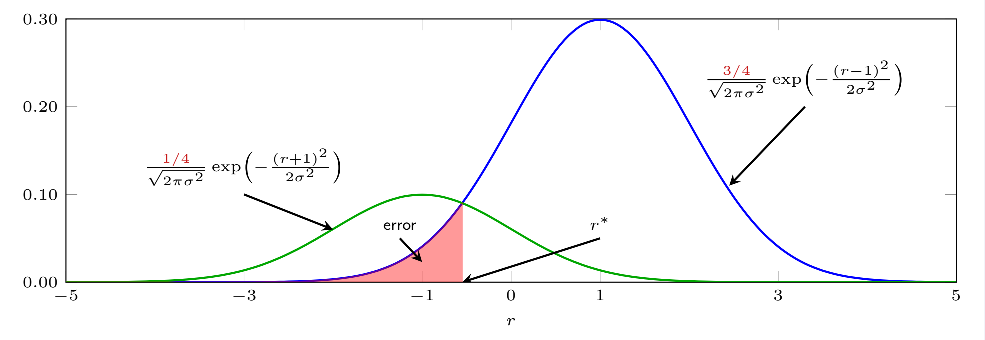
$$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(r-s_1)^2}{2\sigma^2}\right) \geq \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(r-s_2)^2}{2\sigma^2}\right)$$

and $\hat{m}=2$ otherwise.

Example: MAP vs. ML for bi-AGN (1/3)

MAP Thresholds

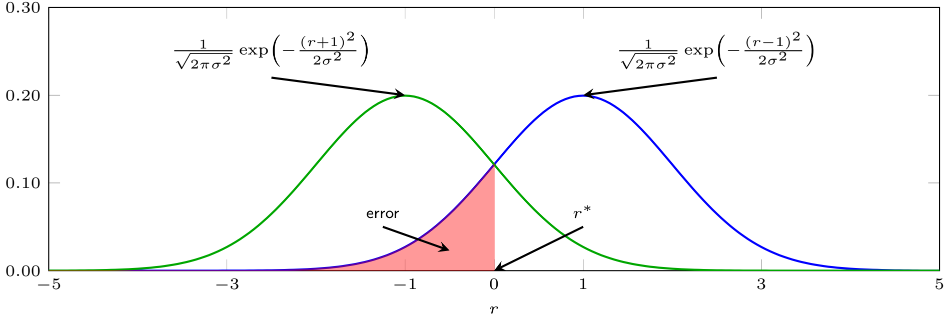
$$\Pr\{M=1\}=3/4,\ s_1=+1, \qquad \Pr\{M=2\}=1/4,\ s_2=-1,$$
$$\implies r^* = -\frac{\ln(3)}{2} \approx -0.549$$



Example: MAP vs. ML for bi-AGN (2/3)

ML Thresholds

$$\Pr\{M=1\}=1/2,\ s_1=+1, \qquad \Pr\{M=2\}=1/2,\ s_2=-1,$$
$$\implies r^* = \frac{1-1}{2} = 0$$



Example: MAP vs. ML for bi-AGN (3/3)

MAP vs. ML

- Threshold in Matlab example was in fact ML detection
- MAP is optimum but more complex than ML
- ML is simple to implement (fixed threshold, one-bit decisions) but suboptimal in general

Summary Module 4.1

Take Home Messages

- In DICO channels the output is continuous (PMFs $\rightarrow$ PDFs)
- AGN model studied in detail
- MAP and ML detectors
  - For bi-AGN, MAP and ML create two intervals via a threshold
  - MAP is optimum but more complex
  - Error probabilities of MAP and ML are not the same

Communication Theory (5ETB0)
Module 4.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

Module 4.2

Presentation Outline

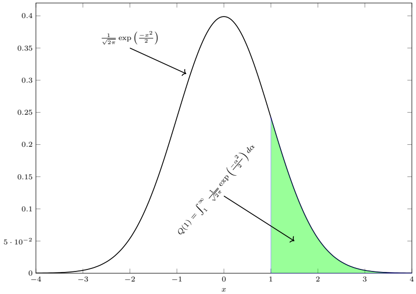
The Q-function

The Q-function

Is a function of $x \in (-\infty, \infty)$:

$$Q(x) \triangleq \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\alpha^2}{2}\right) d\alpha$$

It is the probability that a Gaussian random variable with mean 0 and variance 1 takes a value larger than x.



Useful Property and Three Questions

$$Q(x) = 1 - Q(-x) \implies Q(x) + Q(-x) = 1$$

$$\text{Q1: } Q(0) = ? \quad \text{Q2: } Q(-\infty) = ? \quad \text{and} \quad \text{Q3: } Q(+\infty) = ? \tag{28}$$

Error Probability for bi-AGN Channel

Error Probability for bi-AGN Channel

The error probability of a scalar bi-AGN can be expressed as:

$$
\begin{aligned}
P_{\mathsf{e}} &= \sum_{m\in\mathcal{M}} \Pr\{\hat{M} \neq M | M = m\} \Pr\{M = m\} \\
&= \Pr\{M = 1\}\Pr\{R < r^* | M = 1\} \\
&\qquad + \Pr\{M = 2\}\Pr\{R \geq r^* | M = 2\} \\
&= \Pr\{M = 1\} Q\left(\frac{s_1 - r^*}{\sigma}\right) + \Pr\{M = 2\} Q\left(\frac{r^* - s_2}{\sigma}\right)
\end{aligned}
$$

Error Probability: Derivation

Error Probability for bi-AGN Channel end Equally Likely Messages

Error Probability for Equally Likely Messages

In this case the optimum values is $r^* = 0.5(s_1 + s_2)$ (ML rule), which gives

$$\begin{aligned}
P_\mathrm{e} &= \frac{1}{2}\Pr\{R < r^* | M = 1\} + \frac{1}{2}\Pr\{R \geq r^* | M = 2\} \\
&= \frac{1}{2}Q\left(\frac{s_1 - r^*}{\sigma}\right) + \frac{1}{2}Q\left(\frac{r^* - s_2}{\sigma}\right) \\
&= \frac{1}{2}Q\left(\frac{s_1 - s_2}{2\sigma}\right) + \frac{1}{2}Q\left(\frac{s_1 - s_2}{2\sigma}\right) \\
&= Q\left(\frac{s_1 - s_2}{2\sigma}\right).
\end{aligned}$$

Two Questions

Assume the messages are not equally likely:

- Q1: Can we still use $r^* = 0.5(s_1 + s_2)$?
- Q2: Will the error probability above change?

Example: MAP vs. ML for bi-AGN

Example in Module 4.1

For $s_1 = 1$, $s_2 = -1$, and $\sigma^2 = 1$ with $\Pr\{M = 1\} = 3/4$ and $\Pr\{M = 2\} = 1/4$, the minimum probability of error (achieved by MAP) becomes:

$$\begin{aligned} P_{\mathsf{e}} &= \frac{3}{4} Q\left(1 + \frac{\ln 3}{2}\right) + \frac{1}{4} Q\left(-\frac{\ln 3}{2} + 1\right) \\ &\approx 0.75 \cdot Q(1.5493) + 0.25 \cdot Q(0.4507) \\ &\approx 0.75 \cdot 0.0607 + 0.25 \cdot 0.3261 \\ &\approx 0.1270. \end{aligned}$$

The ML rule would give

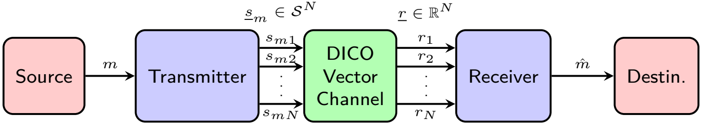
$$P_{\mathsf{e}} = Q\left(\frac{2}{2}\right) \approx 0.1587$$

Does this make sense?

Module 4.2

Presentation Outline

Vector Channels



Definitions

- Source: Produces a *message* $m \in \mathcal{M} \triangleq \{1, 2, \ldots, |\mathcal{M}|\}$ with probability $\Pr\{M = m\}$ for $m \in \mathcal{M}$. The r.v. is M
- Transmitter: Sends a *signal* $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN}) \in \mathcal{S}^N$ if message m is to be transmitted. The random vector is $\underline{S}$
- DICO Vector Channel: Produces output $\underline{r} \in \mathbb{R}^N$ (random vector is $\underline{R}$) with probability density function $p_{\underline{R}}(\underline{r}|\underline{S} = \underline{s}_m) = p_{\underline{R}}(\underline{r}|M = m)$
- Receiver: Forms an *estimate* $\hat{m}$ by observing the received channel output $\underline{r} \in \mathbb{R}^N$ using a mapping $\hat{m} = f(\underline{r}) \in \mathcal{M}$. The r.v. is $\hat{M}$

Decision Variables, MAP and ML

Decision Variables for DICO Vector Channel

The **decision variables** are
$\Pr\{M = m, \underline{R} = \underline{r}\} = \Pr\{M = m\} p_{\underline{R}}(\underline{r}|\underline{S} = \underline{s}_m) = \Pr\{M = m\} p_{\underline{R}}(\underline{r}|M = m)$.

MAP decision rule

The maximum a-posteriori probability (MAP) decision rule is

$$\hat{m}^{\mathsf{MAP}}(\underline{r}) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{M = m | \underline{R} = \underline{r}\} \tag{29}$$

$$= \operatorname*{argmax}_{m \in \mathcal{M}} \Pr\{M = m\} p_{\underline{R}}(\underline{r}|M = m). \tag{30}$$

ML decision rule

The maximum likelihood (ML) decision rule is

$$\hat{m}^{\mathsf{ML}}(\underline{r}) \triangleq \operatorname*{argmax}_{m \in \mathcal{M}} p_{\underline{R}}(\underline{r}|M = m) \tag{31}$$

Module 4.2

Presentation Outline

Part I Error Probability and the Q-function

Part II Vector Channels

Part III Decision Regions

Decision Regions

Decision region for vector channel

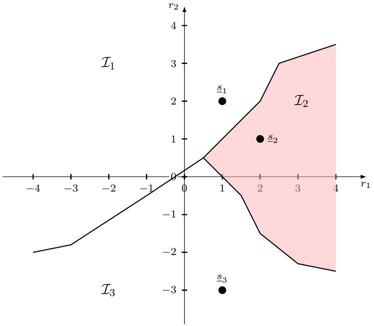
In DICO channels, thresholds define intervals. For DICO vector channels, we need to talk about decision regions

Decision region for vector channel

Given the decision rule $f(\cdot)$ we can write

$$\mathcal{I}_m \triangleq \{\underline{r} \in \mathbb{R}^N : f(\underline{r}) = m\}.$$

$\mathcal{I}_m$ is called the **decision region** that corresponds to message $m \in \mathcal{M}$.



Summary Module 4.2

Take Home Messages

- Q-functions are important to compute error probabilities in the AGN channel
- Detection in vector channels is determinted by *decision regions*
- Error probability in vector channels depend on the channel and detection rule

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Communication Theory (5ETB0)
Module 4.3

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 4.3

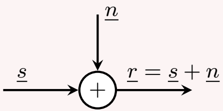
Presentation Outline

AGN Vector Channel

AGN Vector Channel

The AGN vector channel is

$$\underline{r} = \underline{s} + \underline{n},$$



where $\underline{n} \triangleq (n_1, n_2, \ldots, n_N)$ is an N-dimensional noise vector, independent of the signal vector $\underline{S}$, and composed by independent, identically distributed zero-mean Gaussian random variables.

The joint PDF of the noise vector is given by

$$p_{\underline{N}}(\underline{n}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{n_i^2}{2\sigma^2}\right) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N} n_i^2\right)$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\|\underline{n}\|^2}{2\sigma^2}\right)$$

The AGN Vector Channel: A Matlab Example

Conclusions from Example

- Vector AGN noise can be interpreted as multidimensional noise balls
- AGN vector channel can be seen as multiple noise balls centered at $\underline{s}_m$
- Decisions will have to be done in N-dimensional space

Decision Rules for AGN Vector Channel

MAP decision rule for AGN Vector Channel

The conditional PDF for the AGN Vector channel

$$p_{\underline{R}}(\underline{r}|\underline{S}=\underline{s}_m) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\|\underline{r}-\underline{s}_m\|^2}{2\sigma^2}\right)$$

The MAP decision rule is

$$\hat{m}^{\mathsf{MAP}}(\underline{r}) \triangleq \operatorname*{argmin}_{m\in\mathcal{M}} \left\{\|\underline{r}-\underline{s}_m\|^2 - 2\sigma^2 \ln \Pr\{M=m\}\right\}$$

ML decision rule for AGN Vector Channel

The ML decision rule

$$\hat{m}^{\mathsf{ML}}(\underline{r}) \triangleq \operatorname*{argmin}_{m\in\mathcal{M}} \left\{\|\underline{r}-\underline{s}_m\|^2\right\}$$

MAP Derivation

Detailed Derivation

$$
\begin{aligned}
\hat{m}^{\mathsf{MAP}}(\underline{r}) &= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\Pr\{\underline{R}=\underline{r}, \underline{S}=\underline{s}_m\}\right\} = \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\log\Pr\{\underline{R}=\underline{r}, \underline{S}=\underline{s}_m\}\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\log\Pr\{M=m\} p_{\underline{R}}(\underline{r}|\underline{S}=\underline{s}_m)\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\log\Pr\{M=m\}\frac{1}{(2\pi\sigma^2)^{N/2}}\exp\left(-\frac{\|\underline{r}-\underline{s}_m\|^2}{2\sigma^2}\right)\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\log\Pr\{M=m\} - \log(2\pi\sigma^2)^{N/2} - \frac{\|\underline{r}-\underline{s}_m\|^2}{2\sigma^2}\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{\log\Pr\{M=m\} - \frac{\|\underline{r}-\underline{s}_m\|^2}{2\sigma^2}\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmin}} \left\{\frac{\|\underline{r}-\underline{s}_m\|^2}{2\sigma^2} - \log\Pr\{M=m\}\right\} \\
&= \underset{m\in\mathcal{M}}{\operatorname{argmin}} \left\{\|\underline{r}-\underline{s}_m\|^2 - 2\sigma^2\log\Pr\{M=m\}\right\}
\end{aligned}
$$

Decision Rules for AGN Vector Channel

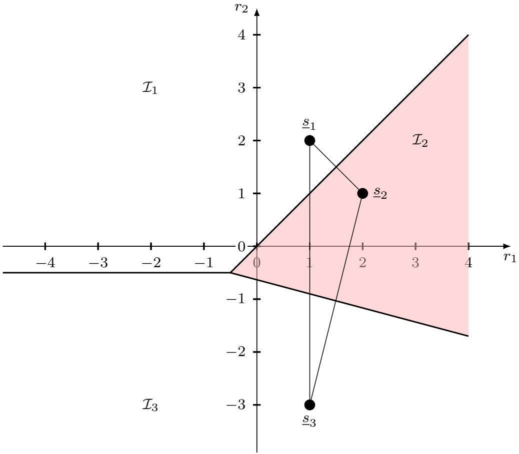
ML decision rule for AGN Vector Channel

- In one dimension (DICO Channel) the optimum threshold was half way between s_1 and s_2
- In N dimensions (DICO Vector Channel) the rule is

$$\hat{m}^{\mathsf{ML}}(\underline{r}) \triangleq \operatorname*{argmin}_{m \in \mathcal{M}} \left\{ \|\underline{r} - \underline{s}_m\|^2 \right\}$$

- For two signals $\underline{s}_1$ and $\underline{s}_2$, this rule corresponds to a hyperplane

ML Decision rule for a 3-signal system



ML Decision Regions: A Matlab Example

Module 4.3

Presentation Outline

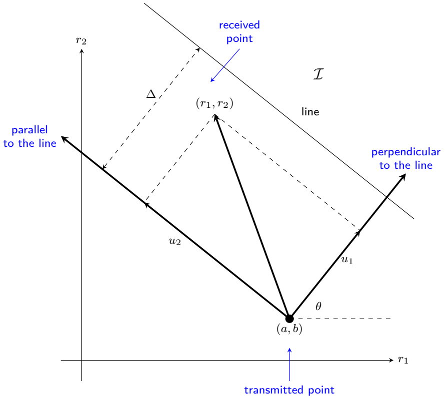
Error Probability: Key Result

AGN vector channel

For the AGN vector channel, the probability that the noise pushes a signal to the wrong side of a hyperplane is

$$P_{\mathcal{I}} = Q\left(\frac{\Delta}{\sigma}\right),$$

where Δ is the distance from the signal-point to the hyperplane and σ^2 is the variance of each noise component.



Error Probability Analysis for ML

Upper Bound on Error Probability (ML)

Average Error Probability: $P_{\mathrm{e}} = \sum_{m \in \mathcal{M}} \Pr\{M = m\} P_{\mathrm{e}}^{m}$

Union bound:

$$P_{\mathrm{e}}^{1} = \Pr\left\{ \bigcup_{m \in \mathcal{M}, m \neq 1} (\|\underline{R} - \underline{s}_m\| \leq \|\underline{R} - \underline{s}_1\|) | M = 1 \right\}$$

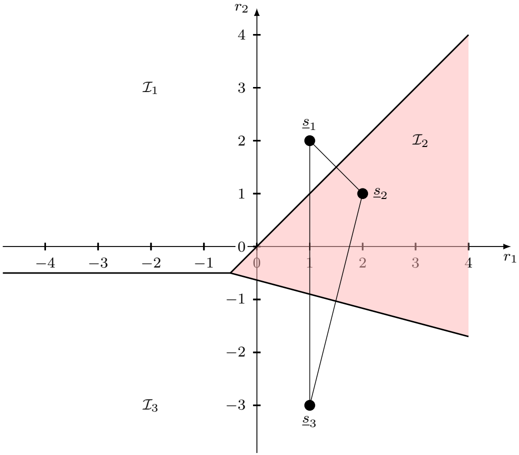
$$\leq \sum_{m \in \mathcal{M}, m \neq 1} \Pr\{\|\underline{R} - \underline{s}_m\| \leq \|\underline{R} - \underline{s}_1\| | M = 1\}$$

$$P_{\mathrm{e}} \leq \sum_{m \in \mathcal{M}} \frac{1}{|\mathcal{M}|} \sum_{m' \in \mathcal{M}, m' \neq m} \Pr\{\|\underline{R} - \underline{s}_{m'}\| \leq \|\underline{R} - \underline{s}_m\| | M = m\}$$

Final Result: AGN channel with per-dimension noise variance σ^2

$$P_{\mathrm{e}} \leq \sum_{m \in \mathcal{M}} \frac{1}{|\mathcal{M}|} \sum_{m' \in \mathcal{M}, m' \neq m} Q\left(\frac{\Delta_{m'm}}{\sigma}\right), \quad \Delta_{m'm} = \frac{\|\underline{s}_{m'} - \underline{s}_m\|}{2}$$

Upper Bound: Geometric Interpretation

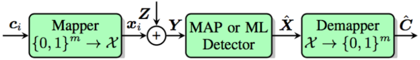


Who Cares?

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 64, NO. 2, FEBRUARY 2018 1231

Asymptotic Comparison of ML and MAP Detectors for Multidimensional Constellations

Alex Alvarado, *Senior Member, IEEE*, Erik Agrell, *Senior Member, IEEE*, and Fredrik Brännström, *Member, IEEE*



II. Preliminaries

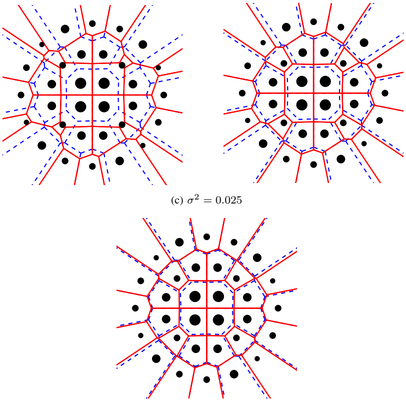
A. System Model

The system model under consideration is shown in Fig. 1. We consider the discrete-time, real-valued, N-dimensional, AWGN channel

$$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{Z}, \tag{1}$$

where the transmitted symbol $\boldsymbol{X}$ belongs to a discrete constellation $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M\}$ and $\boldsymbol{Z}$ is an N-dimensional vector, independent of $\boldsymbol{X}$, whose components are independent and identically distributed Gaussian random variables with zero mean and variance σ^2 per dimension. The conditional channel transition probability is

$$f(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma^2}\right). \tag{2}$$

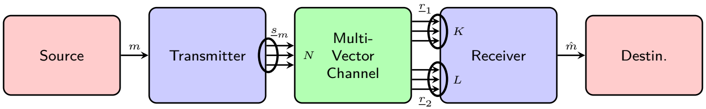


(a) $\sigma^2 = 0.1$ (b) $\sigma^2 = 0.05$ (c) $\sigma^2 = 0.025$

Module 4.3

Presentation Outline

Multi-Vector Channels



Importance

- This model includes for example what is called spatial diversity, i.e., then the transmitter and receiver use multiple antennas (MIMO systems). Used in modern WiFi routers, mobile phones, etc.

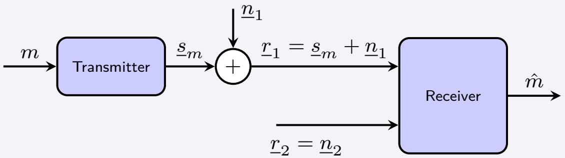
- Theorem of irrelevance: When can we discard $\underline{r}_2$ without affecting performance?

Theorem of Irrelevance

Theorem of Irrelevance (Theorem 4.6)

The output $\underline{r}_2$ of a multi-vector channel is irrelevant (does not affect P_e) if, for all $\underline{r}_1$ and $\underline{r}_2$, the value of $p_{\underline{R}_2}(\underline{r}_2|\underline{S}=\underline{s}_m, \underline{R}_1=\underline{r}_1)$ does not depend on the message m.

Example 4.5

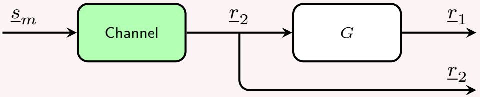


$$p_{\underline{R}_2}(\underline{r}_2|\underline{S}=\underline{s}_m, \underline{R}_1=\underline{r}_1) = p_{\underline{R}_2}(\underline{r}_2) = p_{\underline{N}_2}(\underline{r}_2)$$

Theorem of Reversibility

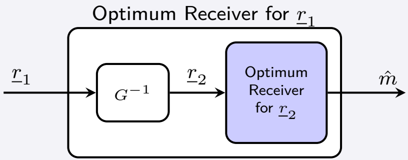
Theorem of Reversibility (Theorem 4.7)

The minimum attainable probability of error is not affected by the introduction of a reversible operation at the output of a channel.



Alternative View

A receiver for $\underline{r}_1$ can be built by first recovering $\underline{r}_2$ from $\underline{r}_1$



Summary Module 4.3

Take Home Messages

- Detection in vector channels is determinted by *decision regions*
- For the AGN vector channel: Euclidean distances!
- Theorems of irrelevance and reversibility let us formally discard certain observations

Communication Theory (5ETB0)
Module 5.1

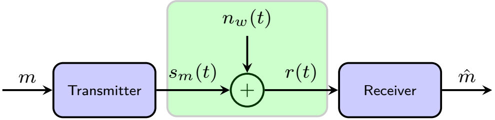
Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

Module 5.1

Presentation Outline

System Description and AWG Noise



Definitions

- Transmitter: Chooses waveform $s_m(t)$ when message m is to be transmitted. Set of used waveforms: $s_1(t), s_2(t), ..., s_{|\mathcal{M}|}(t)$.
- Waveform Channel: Accepts input $s_m(t)$ and adds Gausian noise $n_w(t)$ such that
$$r(t) = s_m(t) + n_w(t)$$
Autocorrelation function of noise process:
$$R_{N_w}(t,s) \triangleq E[N_w(t)N_w(s)] = \frac{N_0}{2}\delta(t-s),$$
- Receiver: Forms an *estimate* $\hat{m}$ based on received waveform $r(t)$.

Module 5.1

Presentation Outline

Part I System Description and AWG Noise

Part II Energy and Orthogonality

Part III Waveform Synthesis

Waveforms: Energy, Orthogonality, and Orthonormality

Energy of a waveform

The energy of a waveform $x(t)$ is defined as

$$E_x \triangleq \int_{-\infty}^{\infty} x^2(t)dt.$$

Orthogonality and orthonormality

The waveforms $\varphi_i(t), i = 1, \dots, N$ are said to be **orthogonal** if

$$\int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt = \begin{cases} E_i & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

If $E_i = 1$ the waveforms are said to be **orthonormal**

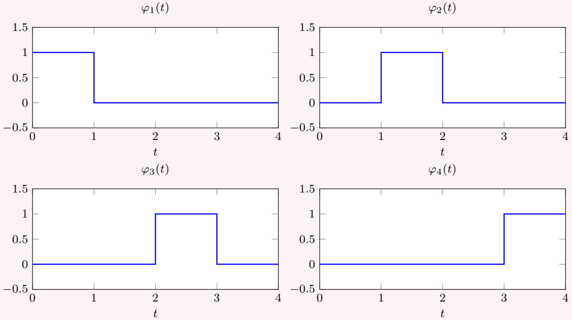
Example 5.1

Time-translated Orthogonal Pulses

Four building-block waveform, time-translated orthogonal pulses:

Question: Are the pulses orthonormal?

$$\int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt = ?$$



Module 5.1

Presentation Outline

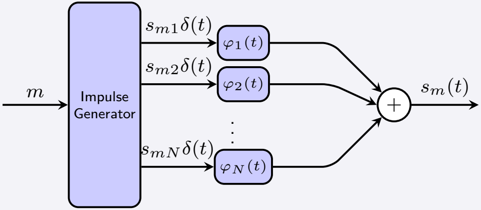
Waveform Synthesis

Waveform Synthesis: From Vectors to Signals

Assume the signal waveform $s_m(t)$ can be expressed as

$$s_m(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t), \text{ for } m \in \mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}.$$

where $\varphi_i(t)$ are called **building-block waveforms**, which are assumed to be *orthonormal*. Signals $s_m(t)$ can be synthesized as:



Canonical transmitter is based on N **building-block waveforms**

Example 5.2

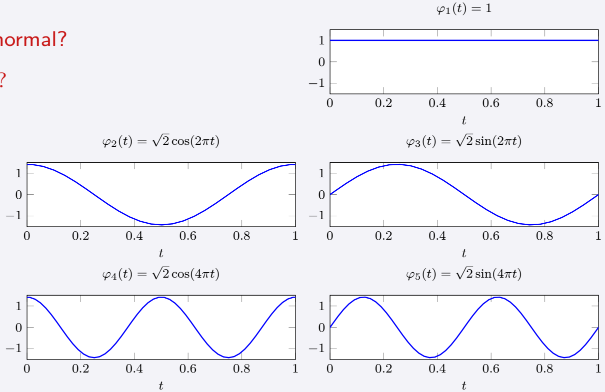
Sine and Cosine Waves

Five building-block waveforms: a pulse with amplitude 1 and four sine and cosine waves. Waveforms are zero outside this time interval.

Question:

Are the pulses orthonormal?

$$\int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt =?$$



Summary Module 5.1

Take Home Messages

- Waveform channels are of great practical importance
- AWGN Channel
- Two signal properties: Energy and Orthogonality/Orthonormality
- From vectors to signals using building-block waveforms

Communication Theory (5ETB0)
Module 5.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 5.2

Presentation Outline

Gram-Schmidt Orthogonalization

Gram-Schmidt Theorem

For an arbitrary signal set, i.e., a set of waveforms $\{s_1(t), s_2(t), \dots, s_{|\mathcal{M}|}(t)\}$ we can construct a set of $N \leq |\mathcal{M}|$ **building-block waveforms** $\{\varphi_1(t), \varphi_2(t), \dots, \varphi_N(t)\}$ and find coefficients s_{mi} such that for $m = 1, 2, \dots, |\mathcal{M}|$ the signals can be synthesized as $s_m(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t)$.

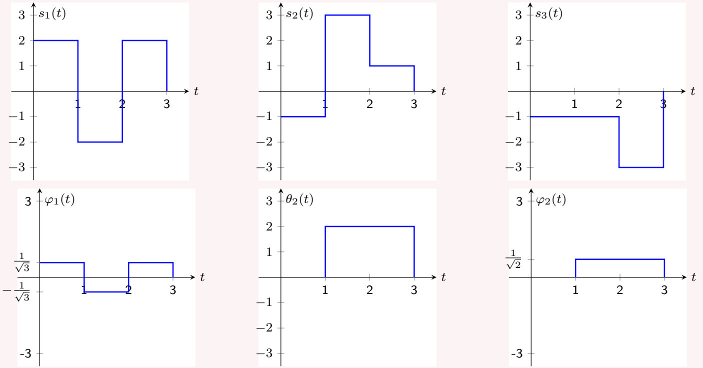
Gram-Schmidt Procedure

1. Compute $\varphi_1(t)$ via $\varphi_1(t) = s_1(t)/\sqrt{E_1}$, then $s_{11} = \sqrt{E_1}$
2. Compute $\theta_2(t) = s_2(t) - s_{21}\varphi_1(t)$, where $s_{21} = \int_{-\infty}^{\infty} s_2(t)\varphi_1(t)dt$
3. Compute $\varphi_2(t) = \theta_2(t)/\sqrt{E_{\theta_2}}$ and $s_{22} = \int_{-\infty}^{\infty} s_2(t)\varphi_2(t)dt$
4. Compute the function $\theta_m(t) \triangleq s_m(t) - \sum_{i=1}^{m-1} s_{mi}\varphi_i(t)$, where the coefficients s_{mi} with $i = 1, \dots, m-1$ are $s_{mi} = \int_{-\infty}^{\infty} s_m(t)\varphi_i(t)dt$
5. Two cases:
   - If $\theta_m(t) \equiv 0$ then we stop, or
   - If $\theta_m(t) \not\equiv 0$ then $\varphi_m(t) = \theta_m(t)/\sqrt{E_{\theta_m}}$ and go back to step 4

Gram-Schmidt Procedure Example

Example 5.3

Three waveforms $s_m(t)$ with $m = 1, 2, 3$



$s_{11} = 2\sqrt{3}$

$(s_{21}, s_{22}) = (-\sqrt{3}, 2\sqrt{2})$

$\theta_3(t) = 0$

$(s_{31}, s_{32}) = (-\sqrt{3}, -2\sqrt{2})$

Module 5.2

Presentation Outline

Part I Gram-Schmidt Orthogonalization

Part II Geometric Interpretation of Signals

Part III Signal Recovery, Irrelevant Data, and Relevant Noise

Geometric Interpretation of Signals (1/2)

Main Points

- We have seen that $s_m(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t)$.
- Thus, to each waveform $s_m(t)$, there corresponds a *vector* with N coefficients $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$.
- *Signal space* is the N-dimensional space where the set of waveforms $\{s_1(t), s_2(t), \ldots, s_{|\mathcal{M}|}(t)\}$ is represented as a set of vectors $\{\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{|\mathcal{M}|}\}$.
- *Signal structure/signal constellation* is the set of vectors $\{\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{|\mathcal{M}|}\}$

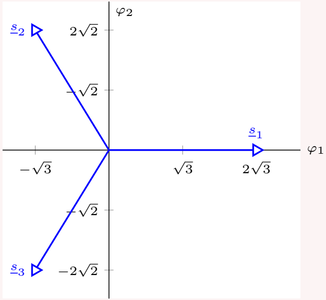
Geomnetric Interpretation of Signals (2/2)

Example 5.3

$$\underline{s}_1 = (s_{11}, s_{12}) = (2\sqrt{3}, 0), \tag{32}$$

$$\underline{s}_2 = (s_{21}, s_{22}) = (-\sqrt{3}, 2\sqrt{2}), \text{ and} \tag{33}$$

$$\underline{s}_3 = (s_{31}, s_{32}) = (-\sqrt{3}, -2\sqrt{2}). \tag{34}$$



Food for Thought

Example 5.3

- "One should also note that a given set of signals can be expanded in many different orthogonal sets of building-block waveforms, possibly with a larger dimension. What remains constant is the geometrical configuration of the vector representations of the signals."
- "It should be noted that, when using the Gram-Schmidt procedure, any ordering of the signals other than $s_1(t), s_2(t), \ldots, s_{|\mathcal{M}|}(t)$ will yield a basis, i.e., a set of building-block waveforms, of the same dimensionality, however in general with different building-block waveforms."

Module 5.2

Presentation Outline

From Signals to Vectors: Without Noise

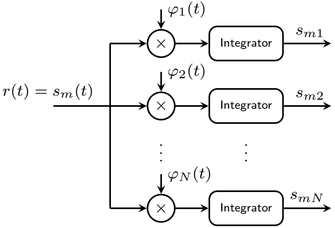
Determine signal vector $\underline{s}_m$ from $r(t) = s_m(t)$

$$\int_{-\infty}^{\infty} s_m(t)\varphi_i(t)dt = \int_{-\infty}^{\infty} \left(\sum_{j=1}^{N} s_{mj}\varphi_j(t) \right) \varphi_i(t)dt$$

$$= \sum_{j=1}^{N} s_{mj} \int_{-\infty}^{\infty} \varphi_j(t)\varphi_i(t)dt = \sum_{j=1}^{N} s_{mj}\delta_{ji} = s_{mi}$$

If we carry this out for all $i = 1, \ldots, N$, we find all coefficients s_{mi} of the vector $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$.

"Note that it is definitely easier to check an N-dimensional vector than a waveform from $-\infty < t < \infty$..."

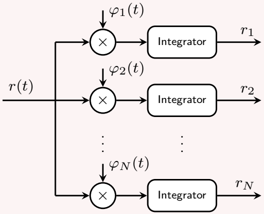


Recovery of Signal Vectors: With Noise

Receiving noisy version $\underline{s}_m$

Receiver gets $r(t) = s_m(t) + n_w(t)$ and computes r_i as

$$r_i \triangleq \int_{-\infty}^{\infty} r(t)\varphi_i(t)dt \text{ for } i = 1, 2, \ldots, N.$$



N-dimensional r-vector is:
$\underline{r}_1 = (r_1, r_2, \ldots, r_N)$

Irrelevant Data

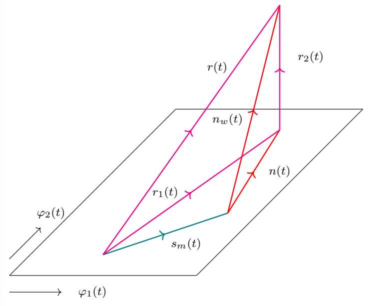
Receiving noisy version $\underline{s}_m$

$r(t) = s_m(t) + n_w(t)$, and thus,

$$\begin{aligned} r_i &= \int_{-\infty}^{\infty} r(t)\varphi_i(t)dt \\ &= \int_{-\infty}^{\infty} s_m(t)\varphi_i(t)dt \\ &\quad + \int_{-\infty}^{\infty} n_w(t)\varphi_i(t)dt \\ &= s_{mi} + n_i \end{aligned}$$

Hence also

$$\underline{r}_1 = \underline{s}_m + \underline{n}$$



with $\underline{r}_1 = (r_1, r_2, \ldots, r_N)$, $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$, and $\underline{n} = (n_1, n_2, \ldots, n_N)$.

Joint Density of Relevant Noise

Irrelevant Noise

Only noise on the signal space is relevant for detection. Noise in all other dimensions can be safely discarded. Proof follows from (i) theorem of reversibility and (ii) theorem of irrelevance.

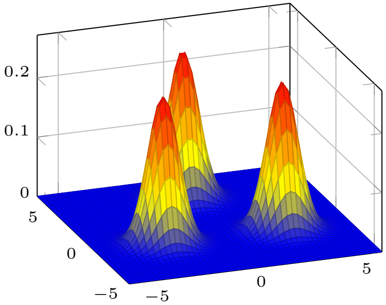
Joint PDF of Relevant Noise

Therefore the joint PDF of the relevant noise vector $\underline{n}$ is

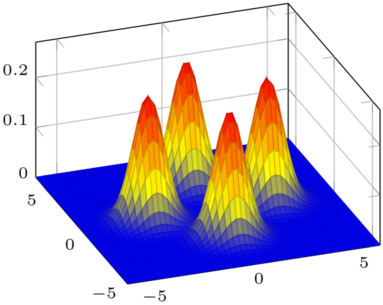
$$p_{\underline{N}}(\underline{n}) = \frac{1}{(\pi N_0)^{N/2}} \exp\left(-\frac{\|\underline{n}\|^2}{N_0}\right),$$

hence the noise is **spherically symmetric** and depends on the magnitude but not on the direction of $\underline{n}$. Noise projected on each direction has variance $\frac{N_0}{2}$.

PDF of 2D ($N = 2$) Signals with $|\mathcal{M}| = 3, 4$



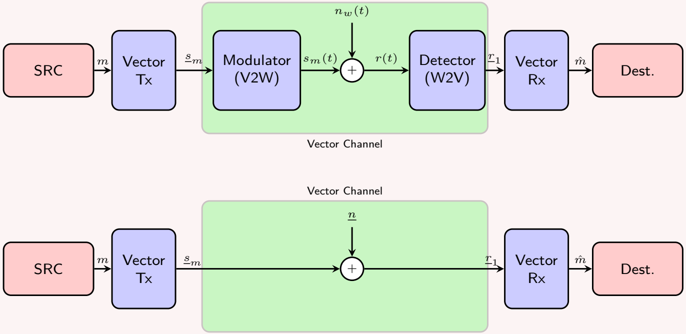
Conditional PDFs for Example 5.3.



Conditional PDFs for Example 5.4.

Relation Between Waveform and Vector Channel

Waveform channel $\Longleftrightarrow$ Vector channel



Summary Module 5.2

Take Home Messages

- Waveform channels are difficult to deal with
- The waveform channel we considered (AWGN) can be converted into a vector DICO channel

Communication Theory (5ETB0)
Module 6.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 6.1

Presentation Outline

Vector Representation of Signals and Operations

Functions $f(t)$ and $g(t)$ and their Vector Representations

Waveforms $f(t)$ and $g(t)$ and orthonormal base $\{\varphi_i(t), i = 1, 2, \cdots, N\}$:

$$f(t) = \sum_{i=1}^{N} f_i \varphi_i(t), \qquad g(t) = \sum_{i=1}^{N} g_i \varphi_i(t)$$

sentations $f(t)$ and $g(t)$:

$$\underline{f} = (f_1, f_2, \cdots, f_N), \qquad \underline{g} = (g_1, g_2, \cdots, g_N)$$

Vector Operations

$$(\underline{f} \cdot \underline{g}) \triangleq \sum_{i=1}^{N} f_i g_i, \quad \|\underline{f}\|^2 \triangleq (\underline{f} \cdot \underline{f}) = \sum_{i=1}^{N} f_i^2$$

$$\|\underline{f} - \underline{g}\|^2 = \|\underline{f}\|^2 + \|\underline{g}\|^2 - 2(\underline{f} \cdot \underline{g})$$

Vector Representation of Signals: Correlation $\Leftrightarrow$ Dot product

Correlation between $f(t)$ and $g(t)$ is the dot product

$$\int_{-\infty}^{\infty} f(t)g(t)dt = (\underline{f} \cdot \underline{g})$$

Proof

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)g(t)dt &= \int_{-\infty}^{\infty} \sum_{i=1}^{N} f_i\varphi_i(t) \sum_{j=1}^{N} g_j\varphi_j(t)dt \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} f_i g_j \int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} f_i g_j \delta_{ij} dt = \sum_{i=1}^{N} f_i g_i = (\underline{f} \cdot \underline{g})\end{aligned}$$

Vector Representation of Signals: Energy $\Leftrightarrow$ Square Norm

Signal Energy

If $g(t) = f(t)$, then the previous result gives

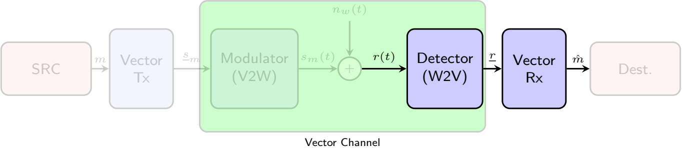
$$\int_{-\infty}^{\infty} f^2(t)dt = (\underline{f} \cdot \underline{f}) = \|\underline{f}\|^2$$

Energy of $f(t)$ is therefore $\|\underline{f}\|^2$.

Energy of $s_m(t)$ is the square norm

$$E_m \triangleq \int_{-\infty}^{\infty} s_m^2(t)dt = \|\underline{s}_m\|^2$$

Optimum Receiver Implementation



Optimum (MAP) Receiver

Computes $\underline{r} = (r_1, r_2, \dots, r_N)$ where

$$r_i = \int_{-\infty}^{\infty} r(t)\varphi_i(t)dt \text{ for } i = 1, 2, \dots, N$$

and solves

$$\min_{m \in \mathcal{M}} \{\|\underline{r} - \underline{s}_m\|^2 - 2\sigma^2 \ln \Pr\{M = m\}\} = \min_{m \in \mathcal{M}} \{\|\underline{r} - \underline{s}_m\|^2 - N_0 \ln \Pr\{M = m\}\}$$

Optimum Receiver requires... A-priori probabilities, Transmitted vectors, Noise variance, and N-dimensional r-values

Optimum Receiver Implementation

Optimum Receiver

The optimum receiver applies the rule

$$\hat{m}^{\mathsf{MAP}}(\underline{r}) = \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\underline{s}_m) + c_m\}$$

where

$$c_m = \frac{N_0}{2}\ln\Pr\{M=m\} - \frac{E_m}{2}$$

and E_m is the energy of $s_m(t)$, for $m \in \mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$.

Proof

$$\begin{aligned}
\hat{m}^{\mathsf{MAP}}(\underline{r}) &= \operatorname*{argmin}_{m\in\mathcal{M}}\{\|\underline{r}-\underline{s}_m\|^2 - N_0\ln\Pr\{M=m\}\}\\
&= \operatorname*{argmin}_{m\in\mathcal{M}}\{\|\underline{r}\|^2 + \|\underline{s}_m\|^2 - 2(\underline{r}\cdot\underline{s}_m) - N_0\ln\Pr\{M=m\}\}\\
&= \operatorname*{argmin}_{m\in\mathcal{M}}\{\|\underline{s}_m\|^2/2 - (\underline{r}\cdot\underline{s}_m) - N_0/2\ln\Pr\{M=m\}\}\\
&= \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\underline{s}_m) + N_0/2\ln\Pr\{M=m\} - \|\underline{s}_m\|^2/2\}
\end{aligned}$$

Module 6.1

Presentation Outline

How to Implement MAP with Filters?

Correlations are dot products

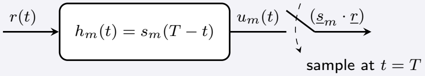
$$\int_{-\infty}^{\infty} r(t)s_m(t)dt = \int_{-\infty}^{\infty} r(t)\sum_{i=1}^{N} s_{mi}\varphi_i(t)dt$$

$$= \sum_{i=1}^{N} s_{mi}\int_{-\infty}^{\infty} r(t)\varphi_i(t)dt = \sum_{i=1}^{N} s_{mi}r_i = (\underline{s}_m\cdot\underline{r})$$

Correlation as Linear Filter and Sampling

$$u_m(t) = \int_{-\infty}^{\infty} r(\alpha)h_m(t-\alpha)d\alpha = \int_{-\infty}^{\infty} r(\alpha)s_m(T-t+\alpha)d\alpha$$

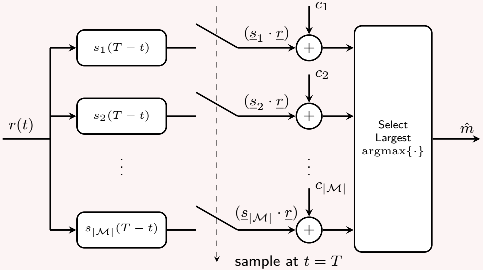
$$\overset{t=T}{=} \int_{-\infty}^{\infty} r(\alpha)s_m(\alpha)d\alpha = (\underline{s}_m\cdot\underline{r})$$



Optimum Receiver: Direct Receiver

Direct Receiver Structure

$$\hat{m}^{\mathsf{MAP}} = \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\underline{s}_m) + c_m\}, \quad c_m = \frac{N_0}{2}\ln \Pr\{M=m\} - \frac{E_m}{2}$$



Two Questions

- Q1: Design a constellation for which the direct receiver is simple
- Q2: How many filters does this receiver need?

Summary Module 6.1

Take Home Messages

- Vectorial representation of signals:
  - Correlations are dot products
  - Energies are square norms
- Correlations as linear filters plus sampling
- Optimum receiver:
  - Finds maximum “shifted” correlations
  - Direct receiver ($|\mathcal{M}|$ Filters)

Communication Theory (5ETB0)
Module 6.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 6.2

Presentation Outline

Motivation for Correlation Receiver

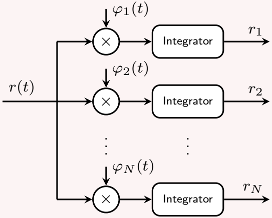
Recovery of Signal Vectors

Gram-Schmidt: $s_m(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t)$

Components r_i are determined as follows:

$$r_i \triangleq \int_{-\infty}^{\infty} r(t)\varphi_i(t)dt$$

for $i = 1, 2, \ldots, N$



Optimum Receiver

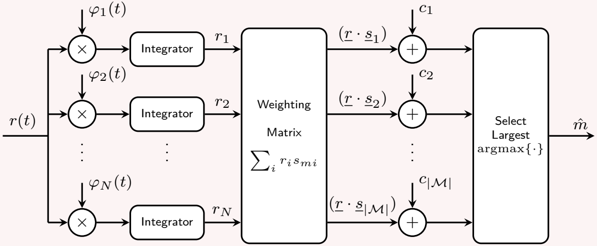
The optimum receiver applies the rule

$$\hat{m}^{\mathsf{MAP}} = \underset{m\in\mathcal{M}}{\operatorname{argmax}}\{(\underline{r}\cdot\underline{s}_m) + c_m\}, \qquad c_m = \frac{N_0}{2}\ln\Pr\{M = m\} - \frac{E_m}{2}$$

Correlation Receiver

Correlation Receiver: Implementation with Matrix Multiplications

Structure:

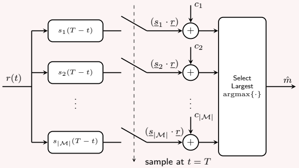


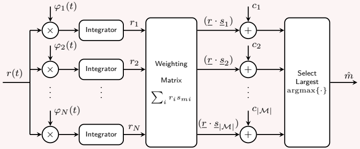
Dot products:

$$\begin{pmatrix} (\underline{r}\cdot\underline{s}_1) \\ (\underline{r}\cdot\underline{s}_2) \\ \vdots \\ (\underline{r}\cdot\underline{s}_{|\mathcal{M}|}) \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1N} \\ s_{21} & s_{22} & \cdots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{|\mathcal{M}|1} & s_{|\mathcal{M}|2} & \cdots & s_{|\mathcal{M}|N} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{pmatrix}$$

Direct Receiver vs. Correlation Receiver

Two Receivers Side-by-side





Three Questions

- Q1: Which receiver is simpler in terms of $\operatorname{argmax}\{\cdot\}$?
- Q2: Which receiver is simpler in terms of filters?
- Q3: Can we always guarantee this?

Example: $s_m(t) = m \cdot p(t)$, $m = 1, 2, 3, \ldots, 128$ and $0 < T < 1$ ps (transmission rate is 7 Gbps). In this case, $N = 1$ and $|\mathcal{M}| = 128 \Rightarrow$ x100 simpler

Module 6.2

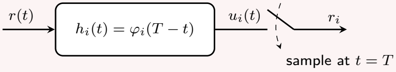
Presentation Outline

Matched-Filter Receiver

Two Important Concepts

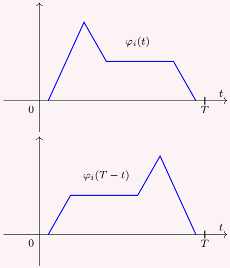
Integral of a multiplication = dot product = Filter+sampling (M6.1)
r-values: Integral of multiplication ($r(t)$ and $\varphi_i(t)$)

Matched Filter Receiver



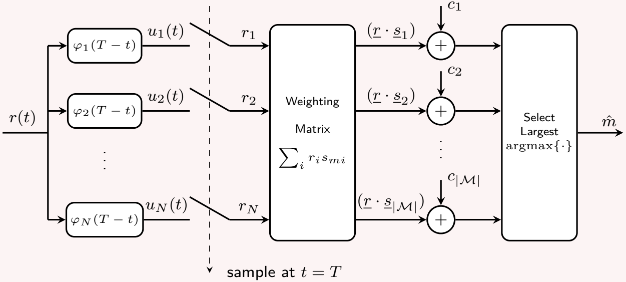
$$u_i(t) = \int_{-\infty}^{\infty} r(\alpha) h_i(t-\alpha) d\alpha$$
$$= \int_{-\infty}^{\infty} r(\alpha) \varphi_i(T-t+\alpha) d\alpha$$
$$\stackrel{t=T}{=} \int_{-\infty}^{\infty} r(\alpha) \varphi_i(\alpha) d\alpha = r_i$$

Building-block $\varphi_i(t)$ and impulse response $\varphi_i(T-t)$:



Matched-Filter Receiver

Matched-filter Receiver Structure



Two Questions

- Q1: Is the matched-filter receiver simpler than the direct receiver?
- Q2: What are differences between the matched-filter and correlation receiver?

Who Cares? We do!

Ecoc 2015 - ID: 0690

Optimum Detection in Presence of Nonlinear Distortions with Memory

Gabriele Liga[(1)], Alex Alvarado[(1)], Erik Agrell[(2)], Marco Secondini[(3)], Robert I. Killey[(1)] and Polina Bayvel[(1)]

(1) Optical Networks Group, University College London, EE Dept., London, UK, g.liga@ee.ucl.ac.uk
(2) Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
(3) TeCIP Institute, Scuola Superiore Sant'Anna, Pisa, Italy

Abstract *The performance of nonlinearity-tailored detection for single channel, single span optical fibre systems is studied. Monotonically decreasing bit error rate with transmitted power can be achieved without any nonlinearity compensation.*

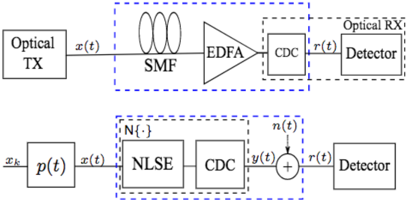


Fig. 1: Single span system and equivalent channel schematic.

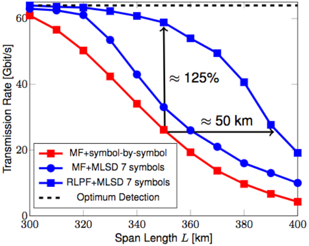


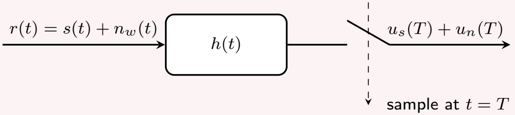
Fig. 4: Achievable transmission rates vs. L for different detection strategies.

Module 6.2

Presentation Outline

Signal and Noise Components

AWGN Through Linear Filter



$$u(T) = \int_{-\infty}^{\infty} r(T-\alpha)h(\alpha)d\alpha = u_s(T) + u_n(T)$$

with

$$\begin{aligned} u_s(T) &\triangleq \int_{-\infty}^{\infty} s(T-\alpha)h(\alpha)d\alpha \\ u_n(T) &\triangleq \int_{-\infty}^{\infty} n_w(T-\alpha)h(\alpha)d\alpha \end{aligned}$$

Signal-to-Noise Ratio (SNR)

SNR Definition

We can now define the *signal-to-noise ratio* as

$$\text{SNR} \triangleq \frac{u_s^2(T)}{E[U_n^2(T)]} \tag{35}$$

Noise variance:

$$\begin{aligned}
E[U_n^2(T)] &= E\left[\int_{-\infty}^{\infty} N_w(T-\alpha)h(\alpha)d\alpha \int_{-\infty}^{\infty} N_w(T-\beta)h(\beta)d\beta\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E[N_w(T-\alpha)N_w(T-\beta)]h(\alpha)h(\beta)d\alpha d\beta \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(\beta-\alpha)h(\alpha)h(\beta)d\alpha d\beta \\
&= \frac{N_0}{2}\int_{-\infty}^{\infty} h^2(\alpha)d\alpha.
\end{aligned}$$

Matched Filter Maximizes the SNR

Schwarz Inequality

For two finite-energy waveforms $a(t)$ and $b(t)$ the inequality

$$\left(\int_{-\infty}^{\infty} a(t)b(t)dt\right)^2 \leq \int_{-\infty}^{\infty} a^2(t)dt \int_{-\infty}^{\infty} b^2(t)dt \tag{36}$$

holds. Equality is obtained only if $b(t) \equiv Ca(t)$ for some constant C.

Result: Maximum Attainable SNR

$$\text{SNR} = \frac{\left[\int_{-\infty}^{\infty} s(T-\alpha)h(\alpha)d\alpha\right]^2}{\frac{N_0}{2}\int_{-\infty}^{\infty} h^2(\alpha)d\alpha} \leq \frac{\int_{-\infty}^{\infty} s^2(T-\alpha)d\alpha \int_{-\infty}^{\infty} h^2(\alpha)d\alpha}{\frac{N_0}{2}\int_{-\infty}^{\infty} h^2(\alpha)d\alpha}$$

$$= \frac{\int_{-\infty}^{\infty} s^2(T-\alpha)d\alpha}{\frac{N_0}{2}} = \frac{E_s}{\frac{N_0}{2}}$$

The Matched-filter Receiver not only minimizes P_{e} but it also maximizes the SNR!

Summary Module 6.2

Take Home Messages

- Correlation receiver
- Matched-filter receiver (correlation receiver using filters)
- Comparison of the three receivers: direct, correlation and matched-filer
- SNR and optimality of Matched-filter receiver

Communication Theory (5ETB0)
Module 7.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 7.1

Presentation Outline

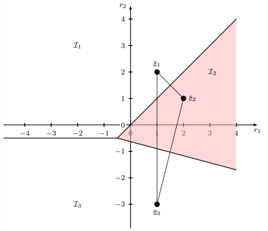
Rotation of Signal Structures

Energy and Average Energy

$$E_{s_m} = \int_0^T s_m^2(t)dt = \|\underline{s}_m\|^2, \qquad E_{\text{av}} \triangleq \sum_{m \in \mathcal{M}} \Pr\{M = m\} E_{s_m} = E[\|\underline{S}\|^2]$$

Error Probability and Energy of Rotated Signals

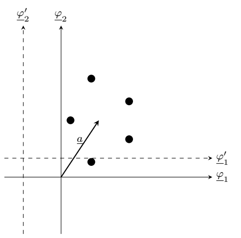
- $\mathcal{S} = \{\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{|\mathcal{M}|}\}$ is *rotated*
- $\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_{|\mathcal{M}|}$ rotated same way
- AWGN vector is spherically symmetric
- P_e will not change
- Av. signal energy will not change

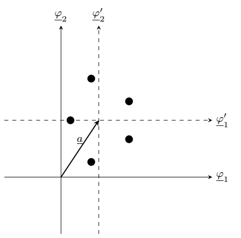


Rotation of Signal Structures: Matlab Example

Translating a Signal Structure

Translating a signal structure does not change P_e



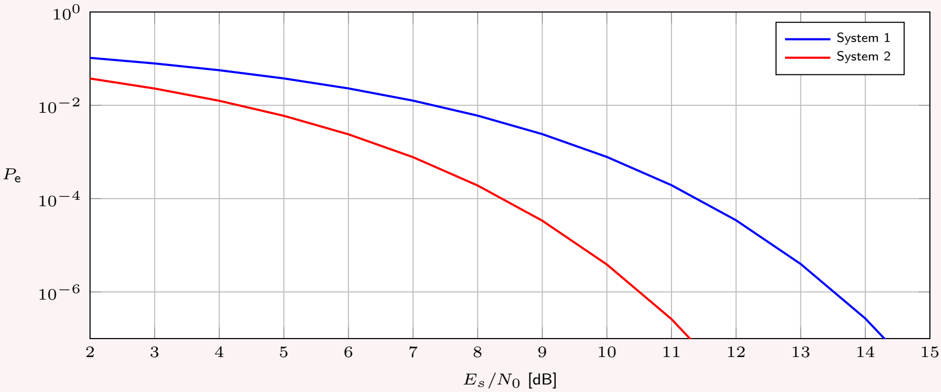


Minimizing the average signal energy

To minimize the average signal energy we should choose the **center of gravity** of the signal structure as the origin of the coordinate system. If the center of gravity of the signal structure $\underline{a} \neq \underline{0}$ we can decrease the average signal energy by $\|\underline{a}\|^2$ by moving the origin of the coordinate system to $\underline{a}$.

Translating a Signal Structure

Translating a signal structure does not change P_e



Module 7.1

Presentation Outline

Binary Orthogonal Signaling

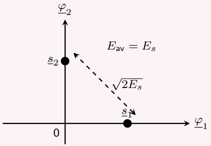
Orthogonal Waveforms (FSK)

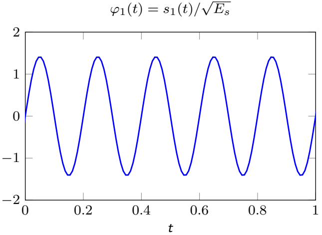
Let $|\mathcal{M}| = \{1, 2\}$ and $\Pr\{M = 1\} = \Pr\{M = 2\} = 1/2$.
Consider two orthogonal waveforms:

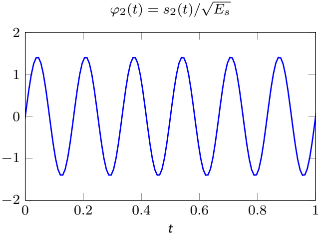
$$s_1(t) = \sqrt{2E_s}\sin(10\pi t), 0 \leq t < 1$$
$$s_2(t) = \sqrt{2E_s}\sin(12\pi t), 0 \leq t < 1$$

Vector representation of signals:

$$\underline{s}_1 = (\sqrt{E_s}, 0),\ \underline{s}_2 = (0, \sqrt{E_s})$$







Module 7.1

Presentation Outline

Binary Antipodal Signaling

Antipodal Waveforms (PSK)

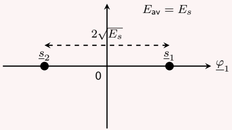
Let $|\mathcal{M}| = \{1,2\}$ and $\Pr\{M=1\} = \Pr\{M=2\} = 1/2$.
Consider two antipodal waveforms:

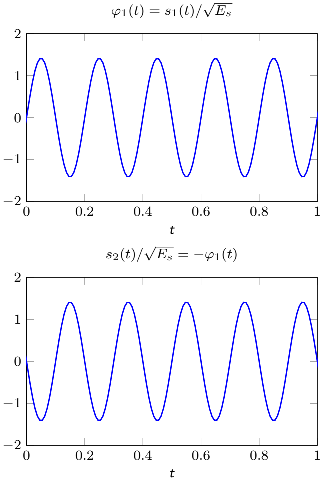
$$s_1(t) = \sqrt{2E_s}\sin(10\pi t), 0 \le t < 1$$

$$s_2(t) = -\sqrt{2E_s}\sin(10\pi t), 0 \le t < 1$$

Vector representation of signals:

$$\underline{s}_1 = (\sqrt{E_s}, 0),\ \underline{s}_2 = (-\sqrt{E_s}, 0)$$





Comparison of Orthogonal and Antipodal Signaling

AGN Vector Channel

For the AGN vector channel, the probability that the noise pushes a signal to the wrong side of a hyperplane is

$$P_{\mathcal{I}} = Q\left(\frac{\Delta}{\sigma}\right),$$

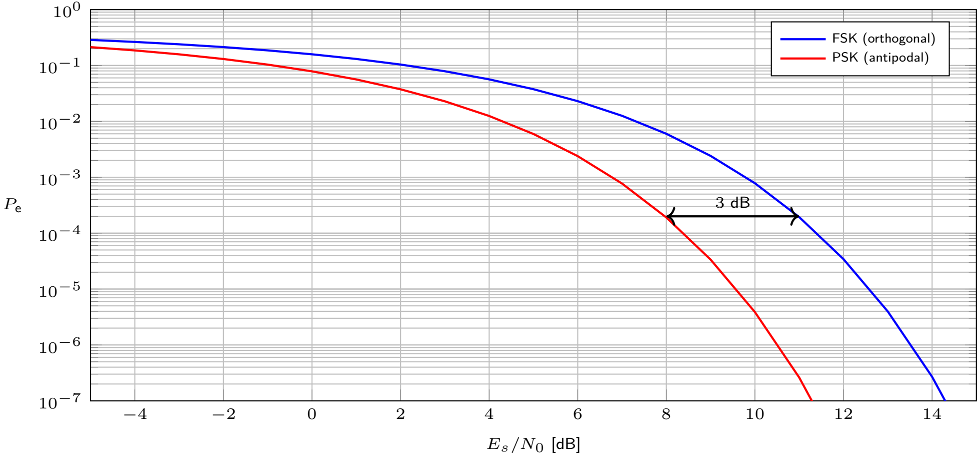
where Δ is the distance from the signal-point to the hyperplane and σ^2 is the variance of each noise component.

Error Probability Comparison

With $E_{av} = E_s$ and power spectral density $N_0/2$, the error probabilities are:

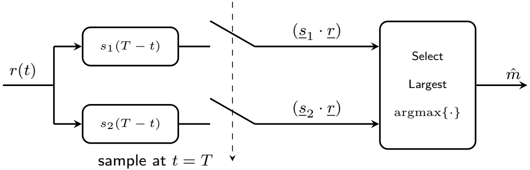
$$P_{\mathrm{e}}^{orth.} = Q\left(\sqrt{E_s/N_0}\right), \quad P_{\mathrm{e}}^{antip.} = Q\left(\sqrt{2E_s/N_0}\right).$$

Comparison of Orthogonal and Antipodal Signaling: Example

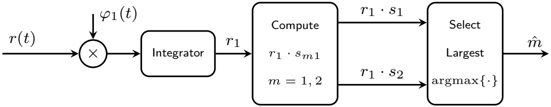


Receivers for Antipodal Signaling

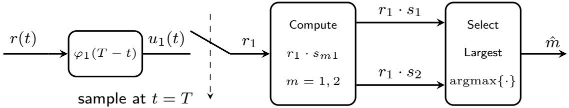
- Direct Receiver



- Correlation Receiver



- Match-filter Receiver



Summary Module 7.1

Take Home Messages

- Rotations do not change the error probability
- Translations save you energy
- Two binary signaling schemes: orthogonal and antipodal.
- Analysis based on building-block waveforms and geometric interpretation of signals

Communication Theory (5ETB0)
Module 7.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 7.2

Presentation Outline

Orthogonal Signal Structures: Definition

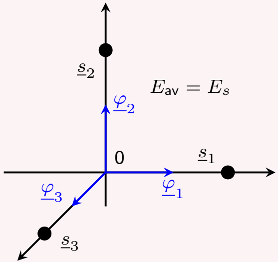
Orthogonal Signal set: Definition

Consider $|\mathcal{M}|$ signals $s_m(t)$ with a-priori probabilities $1/|\mathcal{M}|$ for $m \in \mathcal{M}$. All signals in an orthogonal set are assumed to have equal energy and are orthogonal i.e.,

$$\underline{s}_m \triangleq \sqrt{E_s}\underline{\varphi}_m \text{ for } m \in \mathcal{M},$$

where $\underline{\varphi}_m$ is the unit-vector corresponding to dimension m. There are as many building-block waveforms $\varphi_m(t)$ and dimensions in the signal space as there are messages.

Example: $|\mathcal{M}| = 3$, $\underline{\varphi}_1 = (1,0,0)$, $\underline{\varphi}_2 = (0,1,0)$, $\underline{\varphi}_3 = (0,0,1)$



Orthogonal Signal Structures: Optimum Receiver

Optimum Receiver

Given $\underline{r} = (r_1, r_2, \ldots, r_{|\mathcal{M}|})$, the optimum receiver is:

$$
\begin{aligned}
\hat{m} &= \operatorname*{argmin}_{m\in\mathcal{M}}\{\|\underline{r} - \underline{s}_m\|^2\} \\
&= \operatorname*{argmin}_{m\in\mathcal{M}}\{\|\underline{r}\|^2 + \|\underline{s}_m\|^2 - 2(\underline{r}\cdot\underline{s}_m)\} \\
&= \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\underline{s}_m)\} \\
&= \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\sqrt{E_s}\underline{\varphi}_m)\} \\
&= \operatorname*{argmax}_{m\in\mathcal{M}}\{r_m\}
\end{aligned}
$$

Chose $\hat{m} = i$, where i is the index of the largest component in $\underline{r}$.

Error Probability (1/2)

Model:

$$r_1 = \sqrt{E_s} + n_1, \qquad r_m = n_m, \text{ for } m = 2, 3, \dots, |\mathcal{M}|,$$

$$p_{\underline{N}}(\underline{n}) = \prod_{m=1}^{|\mathcal{M}|} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{n_m^2}{N_0}\right).$$

Correct Prob.:

$$P_{\mathsf{c}} = \int_{-\infty}^{\infty} P_{R_1}(\alpha|M=1) \Pr\{\hat{M}=1|M=1, R_1=\alpha\} d\alpha$$

$$P_{R_1}(\alpha|M=1) = p_N\left(\alpha - \sqrt{E_s}\right) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(\alpha-\sqrt{E_s})^2}{N_0}\right)$$

$$\Pr\{\hat{M}=1|M=1, R_1=\alpha\} = \left(\int_{-\infty}^{\alpha} p_N(\beta) d\beta\right)^{|\mathcal{M}|-1}$$

Error Probability (2/2)

We use $\alpha = \mu\sqrt{N_0/2}$, and thus, the correct probability is

$$P_{\mathsf{c}} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\mu - \sqrt{2E_s/N_0})^2}{2}\right)$$
$$\left(\int_{-\infty}^{\mu\sqrt{N_0/2}} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{\beta^2}{N_0}\right) d\beta\right)^{|\mathcal{M}|-1} d\mu$$

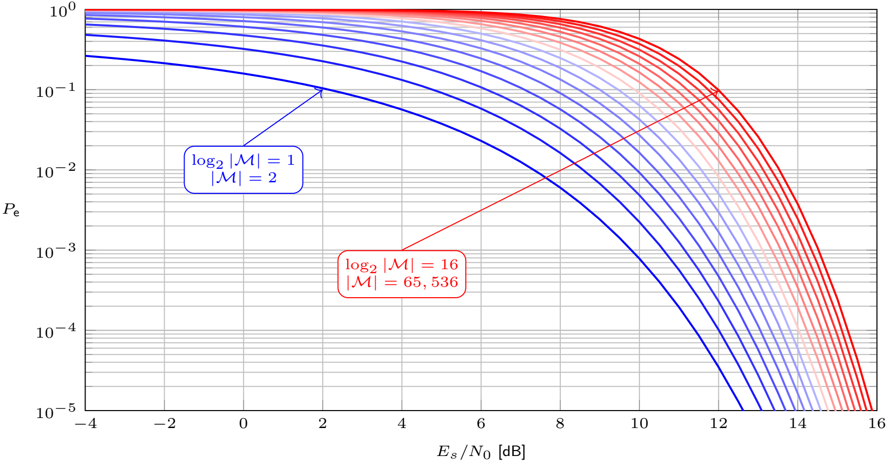
But the inner integral looks familiar... (use $\beta = \lambda\sqrt{N_0/2}$)

$$\int_{-\infty}^{\mu\sqrt{N_0/2}} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{\beta^2}{N_0}\right) d\beta = \int_{-\infty}^{\mu} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\lambda^2}{2}\right) d\lambda$$

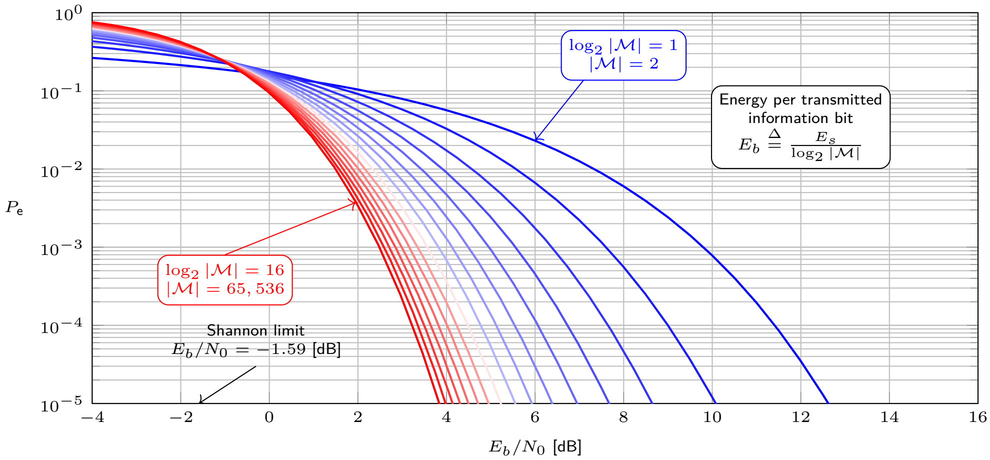
The correct probability is then:

$$P_{\mathsf{c}} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\mu - \sqrt{2E_s/N_0})^2}{2}\right) \big(Q(-\mu)\big)^{|\mathcal{M}|-1} d\mu$$

Error Probability Orthogonal Signals



Error Probability Orthogonal Signals



Module 7.2

Presentation Outline

A Channel Capacity Result

Error Probability and a Capacity Result

The error probability for orthogonal signaling satisfies

$$P_{\mathsf{e}} \leq \begin{cases} 2\exp(-[\sqrt{E_b/N_0} - \sqrt{\ln 2}]^2 \log_2 |\mathcal{M}|) & \ln 2 \leq E_b/N_0 \leq 4\ln 2, \\ 2\exp(-[E_b/(2N_0) - \ln 2] \log_2 |\mathcal{M}|) & 4\ln 2 \leq E_b/N_0. \end{cases}$$

If $E_b > N_0 \ln 2$: (i) Both arguments of the exponentials are negative, and (ii) reliable communication (arbitrarily low error probability) is therefore possible if bit energy is higher than a threshold and $|\mathcal{M}| \to \infty$

Wideband Capacity (Chapter 9)

Reliable transmission of a bit requires at least energy $N_0 \ln 2$, and thus,

$$R = \frac{P_s}{E_b} \leq \frac{P_s}{N_0 \ln 2} \left[\frac{\text{bits}}{\text{seconds}}\right] = C$$

where P_s is the transmitter power.

Energy of Orthogonal Signals

Are Orthogonal Signals Optimal?

Average Energy:

$$E_{\mathsf{av}} = E[\|\underline{S}\|^2] = \sum_{m\in\mathcal{M}} \Pr\{M=m\}\|\underline{s}_m\|^2 = E_s$$

Center of gravity:

$$E\left[\underline{S}\right] = \left(\frac{1}{|\mathcal{M}|}, \frac{1}{|\mathcal{M}|}, \cdots, \frac{1}{|\mathcal{M}|}\right)\sqrt{E_s}$$

But in the limit

$$\lim_{|\mathcal{M}|\to\infty} E\left[\underline{S}\right] = \mathbf{0}$$

Which means:

- Suboptimal in general
- Asymptotically zero loss

Summary Module 7.2

Take Home Messages

- Nonbinary orthogonal signaling: detecion and error probability
- Bit energy vs. symbol energy
- Nonbinary orthogonal signaling leads to a capacity result

Communication Theory (5ETB0)
Module 8.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

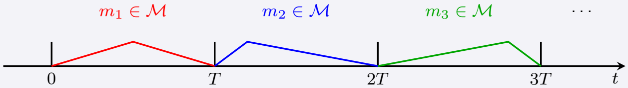
Module 8.1

Presentation Outline

Motivation: A Stream of Messages

Preliminaries

- Previously: considered transmission of a **single randomly-chosen message** $m \in \mathcal{M}$ over a waveform channel
- Now: Transmission of a **stream** of messages over the AWGN waveform channel
- Assumption 1: The signals $s_m(t)$ are only non-zero inside the time-interval $0 \leq t < T$
- Assumption 2: Equally likely messages (i.e., $\Pr\{M = m\} = 1/|\mathcal{M}|$ for all $m \in \mathcal{M}$)



Definitions and Problem Statement

Definitions

- **Transmit Power** is P_s ([Joule/sec] or [Watt])
- **Average Energy** is $E_s = P_s T$ [Joule]
- **Transmission rate** R is defined as

$$R \triangleq \frac{\log_2 |\mathcal{M}|}{T} \quad \left[\frac{\text{bits}}{\text{second}}\right]$$

- **Energy per transmitted bit** is

$$E_b \triangleq \frac{E_s}{\log_2 |\mathcal{M}|} = \frac{E_s}{T} \frac{T}{\log_2 |\mathcal{M}|} = \frac{P_s}{R} \quad \left[\frac{\text{Joule}}{\text{bit}}\right]$$

Questions to be Answered

- What is the **maximum rate at which we can communicate reliably** over a waveform channel when the available power is P_s?
- What are the signals that are to be used to achieve this maximum rate?
- Two systems considered: bit-by-bit and block-orthogonal signaling

Module 8.1

Presentation Outline

Bit-by-Bit Signaling: Definitions

Rate and Transmitted Waveform

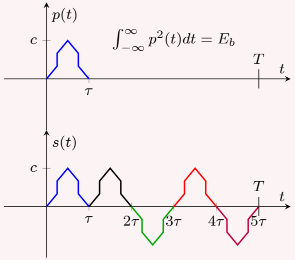
Transmit K binary digits $b_1 b_2 \cdots b_K$ in T seconds. Then

$$|\mathcal{M}| = 2^K, \quad R = \frac{\log_2 |\mathcal{M}|}{T} = \frac{K}{T}$$

Transmit signal $s(t)$, composed of K pulses $p(t)$ that are time shifted:

$$s(t) = \sum_{i=1}^{K} (-1)^{b_i+1} p(t-(i-1)\tau)$$

Signal set is: $\mathcal{S} = \{s_1(t), s_2(t), \ldots, s_{2^K}(t)\}$



Message to be transmitted: 11010

Bit-by-Bit Signaling: Building-block Waveform

Building-block Waveforms

The building-block waveforms are time-shifts over multiples of τ of the normalized pulse $p(t)/\sqrt{E_b}$

$$\varphi_i(t) \triangleq \frac{p(t-(i-1)\tau)}{\sqrt{E_b}}, \qquad i = 1, 2, \ldots, K$$

Questions...

Q1: Can the messages $s_m(t)$ be written as a linear combination of $\varphi_i(t)$? Yes!

$$s(t) = \sum_{i=1}^{K} (-1)^{b_i+1} p(t-(i-1)\tau) = \sum_{i=1}^{K} (-1)^{b_i+1} \sqrt{E_b}\varphi_i(t)$$

Q2: Are $\varphi_i(t)$ orthonormal? Yes!

$$\int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt = \frac{1}{E_b}\int_{-\infty}^{\infty} p(t-(i-1)\tau)p(t-(j-1)\tau)dt = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Q3: What is the dimensionality of the signal space? $N = K$

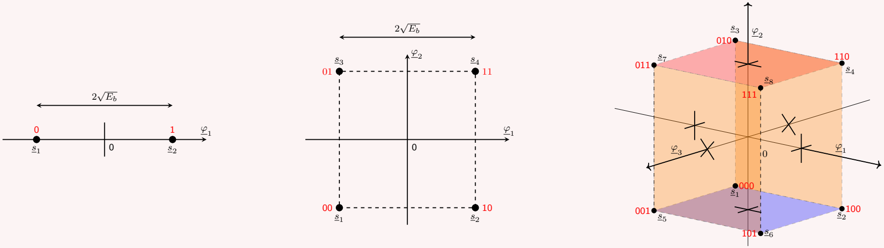
Bit-by-Bit Signaling: Geometry

Geometric Representation

The signals are

$$s(t) = \sum_{i=1}^{K} (-1)^{b_i+1} \sqrt{E_b} \varphi_i(t)$$

The vectorial representation is $\underline{s}_m = \sqrt{E_b}((-1)^{b_1+1}, (-1)^{b_2+1}, \ldots, (-1)^{b_N+1})$



$K = N = 1, |\mathcal{M}| = 2$ $K = N = 2, |\mathcal{M}| = 4$ $K = N = 3, |\mathcal{M}| = 8$

Bit-by-Bit Signaling: Reception

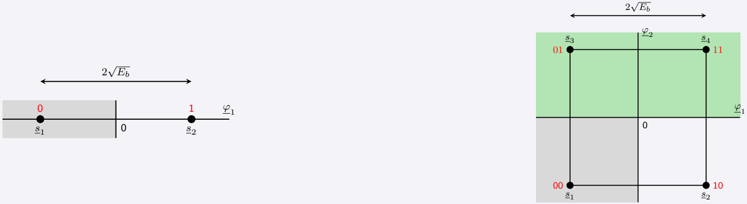
Optimum Receiver

The optimum receiver decides $\hat{m} = 1$ if

$$r_i < 0, \text{ for all } i = 1, \dots, K$$

Geometric Interpretation

If $m = 1$ is transmitted: $\underline{s}_1 = \left(-\sqrt{E_b}, -\sqrt{E_b} \dots, -\sqrt{E_b}\right)$



Note: To estimate b_i with $i = 1, 2, \dots, K$, only r_i in dimension i is needed.

Bit-by-Bit Signaling: Error Probability (1/2)

Correct and Error Probabilities

- The signal hypercube is symmetrical
- Assume that $\underline{s}_1$ was transmitted
- No error occurs if $r_i = -\sqrt{E_b} + n_i < 0$ for all $i = 1, \dots, K$:

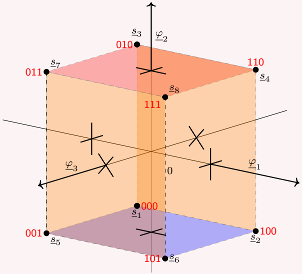
$$n_i < \sqrt{E_b} \text{ for all } i = 1, \dots, K$$

- Correct Probability

$$P_\mathsf{c} = \left(1 - Q(\sqrt{2E_b/N_0})\right)^K$$

- Error Probability

$$P_\mathsf{e} = 1 - \left(1 - Q(\sqrt{2E_b/N_0})\right)^K$$



Bit-by-Bit Signaling: Error Probability (2/2)

Error Probabilities Considerations

- Using $K = RT$ and $E_b = P_s/R$

$$P_{\mathsf{e}} = 1 - \left(1 - Q\left(\sqrt{\frac{2P_s}{RN_0}}\right)\right)^{RT}$$

- Fix P_s and R and consider two extreme cases for T:
  - $T = 1/R \Rightarrow K = 1 \Rightarrow$

$$P_{\mathsf{e}} = Q\left(\sqrt{\frac{2P_s}{RN_0}}\right)$$

    Conclusion: P_{e} can be decreased by increasing P_s or by decreasing R
  - $T \to \infty \Rightarrow P_{\mathsf{e}} \to 1$
    Conclusion: Reliability **cannot** be increased by increasing T

Is this the end of the story?

Can we increase reliability by increasing T? Yes! With block-orthogonal signaling

Summary Module 8.1

Take Home Messages

- Introduced the problem of serial transmission
- Bit-by-bit signalling model and analysis
- Increasing dimensionality in bit-by-bit signalling does not help

Communication Theory (5ETB0)
Module 8.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 8.2

Presentation Outline

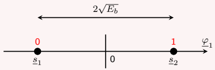
Recap: Bit-by-Bit Signaling

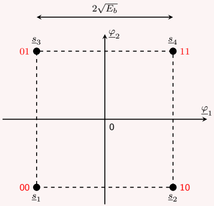
Bit-by-Bit Signaling

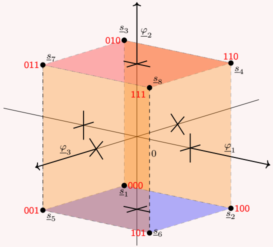
The signals are

$$s(t) = \sum_{i=1}^{K} (-1)^{b_i+1} \sqrt{E_b} \varphi_i(t)$$

The vectorial representation is $\underline{s}_m = \sqrt{E_b}((-1)^{b_1+1}, (-1)^{b_2+1}, \ldots, (-1)^{b_N+1})$







Conclusion: Reliability **cannot** be increased by increasing T

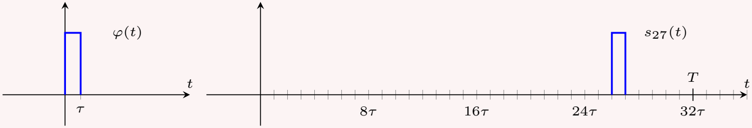
Block-Orthogonal Signaling: Description

System Description

- Transmit K bits in T seconds. Rate is $R = K/T$
- Send one out of 2^K orthogonal pulses every T seconds
- Focus on pulse-position modulation (PPM) where $|\mathcal{M}| = 2^K$ signals are

$$s_m(t) = \sqrt{E_s}\varphi(t-(m-1)\tau), \text{ for } m = 1, ..., 2^K,$$

  with $\varphi(t)$ a unit-energy pulse and duration τ less than $T/2^K$
- All signals *within the block* $[0, T)$ are orthogonal and have energy E_s.



$\int_{-\infty}^{\infty} \varphi^2(t)dt = 1$ $\qquad$ $\int_{-\infty}^{\infty} s_{27}^2(t)dt = E_s = TP_s$

Block-Orthogonal Signaling: Error Probability

Error Probabilities Considerations

We assume that the energy per transmitted bit of information is

$$E_b/N_0 = (1+\epsilon)^2 \ln 2,$$

with $0 \leq \epsilon \leq 1 \Rightarrow$ We are willing to spend slightly more than $N_0 \ln 2$ [Joule]
Using $\log_2 |\mathcal{M}| = RT$ in the upper bound in P_e for $E_b/N_0 \geq \ln 2$:

$$\begin{aligned} P_\mathsf{e} &\leq 2\exp(-[\sqrt{E_b/N_0} - \sqrt{\ln 2}]^2 \log_2 |\mathcal{M}|) \\ &= 2\exp(-[\sqrt{(1+\epsilon)^2 \ln 2} - \sqrt{\ln 2}]^2 RT) = 2\exp(-\epsilon^2 RT \ln 2) \end{aligned}$$

We now use $E_b = P_s/R$ to get

$$\boxed{R = \frac{1}{(1+\epsilon)^2} \frac{P_s}{N_0 \ln 2}, \qquad P_\mathsf{e} \leq 2\exp(-\epsilon^2 RT \ln 2)}$$

What happens if we change ϵ and T?

Block-Orthogonal Signaling: Capacity Result

Channel Capacity

With available average power P_s we can achieve rates R smaller than but arbitrarily close to

$$C_\infty \triangleq \frac{P_s}{N_0 \ln 2} \quad \left[\frac{\text{bit}}{\text{second}}\right]$$

while the error probability P_e can be made arbitrarily small by increasing T.

Extra Comments

- The reliability can be increased not only by increasing the power P_s or decreasing the rate R (bit-by-bit signalling) but also by increasing the "codeword-lengths" T.
- The channel **capacity** C_∞ depends only on the available power P_s and power spectral density $N_0/2$ of the noise.
- Only rates up to the capacity can be achieved (not possible to go beyond)
- Warning: Is this the end of the story? No. We have ignored the dimensionality of the signal sets...

Dimensions Needed for Bit-by-Bit and for Block-Orthogonal Signaling

Bit-by-Bit and Block-Orthogonal

Signaling	Dimensions per block	Dimensions per second
Bit-by-Bit	$K = RT$	$K/T = R$
Block Orthogonal	$2^K = 2^{RT}$	$2^K/T = 2^{RT}/T$

Block-orthogonal signaling: Pros and Cons

For a given rate R:

- Arbitrarily high reliability can be achieved by increasing T (not the case for bit-by-bit signaling)
- The number of dimensions per second explodes by increasing T
- Bad because a channel with a finite bandwidth cannot accommodate all these dimensions

Module 8.2

Presentation Outline

Part I Block-Orthogonal Signaling

Part II Dimensions and Bandwidth

Bandwidth, time, and dimensions

The Dimensionality Theorem

Let $\varphi_i(t)$, for $i = 1, ..., N$ denote any set of orthonormal waveforms. Assume that for all waveforms $\varphi_i(t)$ for $i = 1, \ldots, N$

- $\varphi_i(t) = 0$ for all t outside $[0, T)$, and
- its Fourier transform satisfies $\int_{-W}^{+W} |\Phi_i(f)|^2 df \approx 1$.

Then the number of orthogonal waveforms (dimensions) N is (roughly) upper-bounded by $2WT$. The parameter W is called **bandwidth** (in Hz).

Comments

- Number of waveforms cannot be much more than approximately $2WT$
- The number of dimensions per second is not much more than $2W$
- It can be shown that instead of $2^K/T$, we could have $(2K+1)/T$ dimensions per second.

Summary Module 8.2

Take Home Messages

- Block orthogonal sinaling leads to a capacity result
- The number of dimensions per second explodes
- Dimensionality theorem tell us how good/bad this is w.r.t. bandwidth

Communication Theory (5ETB0)
Module 9.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 9.1

Presentation Outline

Motivation and Objective

Motivation

- **Dimensionality Theorem**: $\approx 2W$ dimensions per second
- **Block-orthogonality signaling:**
  - Requires a lot of dimensions per second
  - Can give $P_{\mathrm{e}} \to 0$
- **Bit-by-bit signaling:**
  - Requires as many dimensions as there are bits to be transmitted
  - Can only made reliable ($P_{\mathrm{e}} \to 0$) by increasing the transmitter power P_s or by decreasing the rate R

Module Objective

Can we achieve reliable transmission ($P_{\mathrm{e}} \to 0$) at certain rate R by increasing T, when both the bandwidth W and available power P_s are fixed?

Module 9.1

Presentation Outline

Channel Capacity of AGN Vector Channel (1/2)

Capacity of AGN Vector Channel

For the AGN vectorial channel, there exist for N large enough, sets of $|\mathcal{M}|$ vectors, $\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{|\mathcal{M}|}$ where $\|\underline{s}_m\|^2 \approx N E_N$ for all $m = 1, 2, \ldots, |\mathcal{M}|$ and where $P_{\mathrm{e}} \approx 0$ as long as

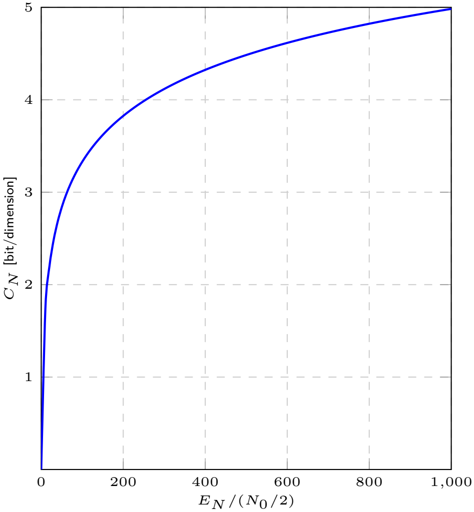
$$R_N = \frac{\log_2 |\mathcal{M}|}{N} < C_N \triangleq \frac{1}{2} \log_2 \left(\frac{E_N + N_0/2}{N_0/2} \right) \quad \left[\frac{\text{bit}}{\text{dimension}} \right]$$

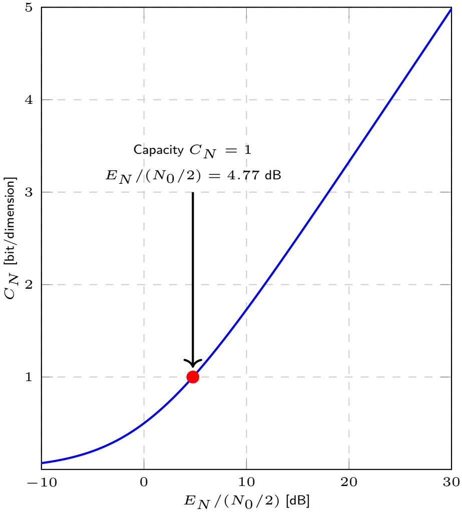
where all vectors have length N, are equiprobable, and where E_N is the available energy per dimension.

Comments

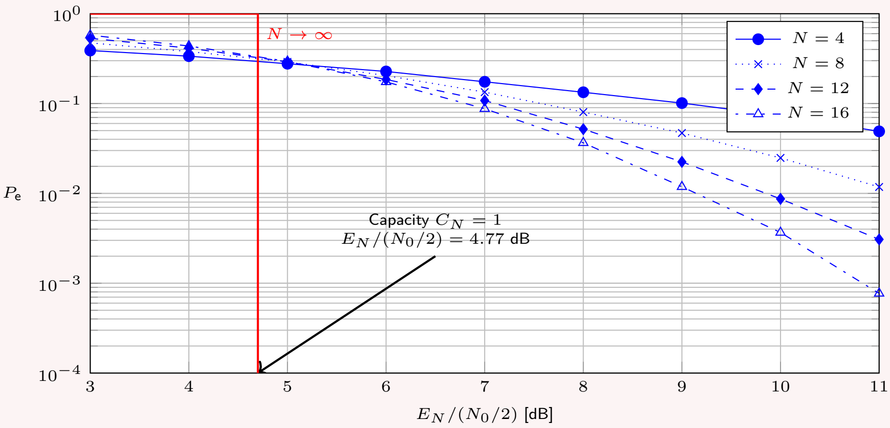
- Fundamental limit in terms of rate
- Elements in the vectors (codewords) are i.i.d. Gaussian random variables
- Vectors (codewords) are on the shell of a multidimensional sphere
- Converse: It can also be shown that P_{e} cannot be small if the rate per dimension $R_N > C_N$

Channel Capacity of AGN Vector Channel (2/2)





Example: Randomly Generated Codes



Module 9.1

Presentation Outline

Channel Capacity of Baseband AWGN Channel

Capacity of Baseband AWGN Channel

For a waveform channel with spectral noise density $N_0/2$, frequency bandwidth W, and available transmitter power P_s, the capacity in bit per second is

$$C \triangleq W \log_2 \left(1 + \frac{P_s}{W N_0}\right) \left[\frac{\text{bit}}{\text{second}}\right].$$

Thus reliable communication $(P_\text{e} \to 0)$ is possible for rates R in bit per second smaller than C, while rates larger than C are not realizable with $P_\text{e} \to 0$.

Questions and Comments

Q1: What is more benefitial: Bandwidth or power?
Answer: Linear growth with bandwidth. Use bandwidth when available
C1: The argument of the logarithm is $1 +$ SNR
C2: We will study two extreme cases: SNR $\ll 1$ and SNR $\gg 1$

Vectorial AGN vs. Baseband AWGN Channels

Are they the same?

$$C = W\log_2\left(1+\frac{P_s}{WN_0}\right)\ \left[\frac{\text{bit}}{\text{second}}\right],\ \text{vs. } C_N = \frac{1}{2}\log_2\left(\frac{E_N+N_0/2}{N_0/2}\right)\ \left[\frac{\text{bit}}{\text{dimension}}\right]$$

$$\begin{aligned}\frac{C}{2W} &= \frac{1}{2}\log_2\left(1+\frac{P_s}{WN_0}\right)\ \left[\frac{\text{bit}}{\text{dimension}}\right]\\ &= \frac{1}{2}\log_2\left(1+\frac{E_s}{TWN_0}\right)\ \left[\frac{\text{bit}}{\text{dimension}}\right]\end{aligned}$$

But $E_N = E_s/N$ and $T \approx N/(2W)$, and thus,

$$\frac{E_s}{TWN_0} = \frac{NE_N}{N/2N_0}$$

$$\frac{C}{2W} = \frac{1}{2}\log_2\left(1+\frac{E_N}{N_0/2}\right)\ \left[\frac{\text{bit}}{\text{dimension}}\right] = \frac{1}{2}\log_2\left(\frac{N_0/2+E_N}{N_0/2}\right)\ \left[\frac{\text{bit}}{\text{dimension}}\right]$$

Channel Capacity of the Wideband AWGN Channel

Capacity of the Wideband AWGN Channel (Power-Limited)

The capacity C_∞ of the **wideband AWGN channel** with power spectral density $N_0/2$, when the transmitter power is P_s, is given by

$$C_\infty = \frac{P_s}{N_0 \ln 2}$$

Derivation of Wideband AWGN Channel Capacity

$$\begin{aligned}\lim_{W\to\infty} W \log_2\left(1+\frac{P_s}{WN_0}\right) &= \lim_{W\to\infty} \frac{\log_2\left(1+\frac{P_s}{WN_0}\right)}{1/W}\\ &= \lim_{W\to\infty} \frac{P_s}{N_0} \frac{-1/W^2}{\left(1+\frac{P_s}{WN_0}\right)\ln(2)} \frac{1}{-1/W^2}\end{aligned}$$

where we used $\frac{d}{dx}\log_2(x) = \frac{1}{x\ln(2)}$.

Relations between Capacities and SNR

Capacity of AWGN Channel at High-SNRs (Bandwidth-Limited)

The capacity of the AWGN is $C = W\log_2(1+\mathsf{SNR})$, where where $\mathsf{SNR} \triangleq P_s/(WN_0)$. When $\mathsf{SNR} \gg 1$, the capacity can be approximated as $W\log_2(\mathsf{SNR})$.

Power-limited and Bandwidth-limited Regimes

We can can distinguish between two cases.

$$C \approx \begin{cases} P_s/(N_0 \ln 2) & \text{if SNR} \ll 1, \\ W\log_2(\mathsf{SNR}) & \text{if SNR} \gg 1 \end{cases}$$

- The case $\mathsf{SNR} \ll 1$ is called the **power-limited** regime. There is enough bandwidth.
- When $\mathsf{SNR} \gg 1$ we speak about **bandwidth-limited** channels.

Summary Module 9.1

Take Home Messages

- Capacity of the AGN vector channel
- Capacity of the baseband AWGN channel
- Power-limited and bandwidth-limited regimes
- Performance of random codes

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Communication Theory (5ETB0)
Module 9.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

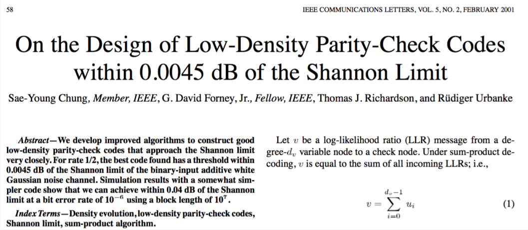
Module 9.2

Presentation Outline

How do we achieve capacity?

Modern Codes

- Turbo Codes, invented in 1990-1991, “to good to be true”. Used in 3G and 4G standards. [1]
- Low-density parity check codes, invented in 1960 and re-discovered in 1996. Used in WiFi and DVB.



58 IEEE COMMUNICATIONS LETTERS, VOL. 5, NO. 2, FEBRUARY 2001

On the Design of Low-Density Parity-Check Codes within 0.0045 dB of the Shannon Limit

Sae-Young Chung, *Member, IEEE*, G. David Forney, Jr., *Fellow, IEEE*, Thomas J. Richardson, and Rüdiger Urbanke

***Abstract*—We develop improved algorithms to construct good low-density parity-check codes that approach the Shannon limit very closely. For rate 1/2, the best code found has a threshold within 0.0045 dB of the Shannon limit of the binary-input additive white Gaussian noise channel. Simulation results with a somewhat simpler code show that we can achieve within 0.04 dB of the Shannon limit at a bit error rate of 10^{-6} using a block length of 10^7.**

***Index Terms*—Density evolution, low-density parity-check codes, Shannon limit, sum-product algorithm.**

Let v be a log-likelihood ratio (LLR) message from a degree-d_v variable node to a check node. Under sum-product decoding, v is equal to the sum of all incoming LLRs; i.e.,

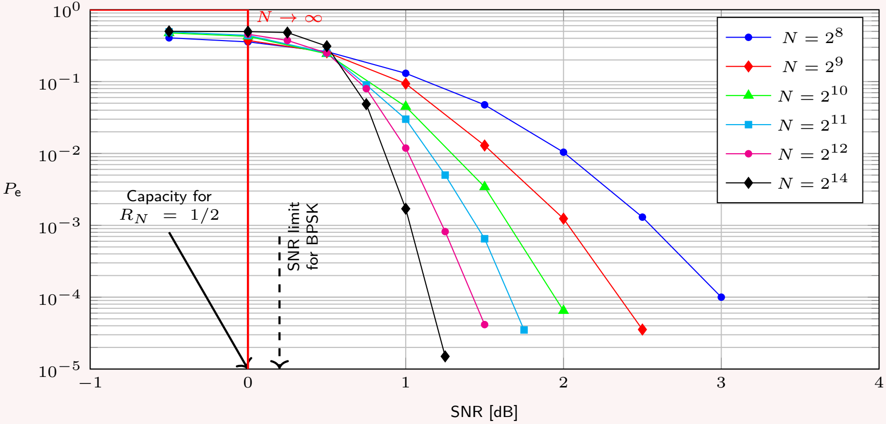
$$v = \sum_{i=0}^{d_v-1} u_i \qquad (1)$$

- Polar Codes, first codes with explicit construction that can be proven to achieve the channel capacity. Used in 5G NR. [2]

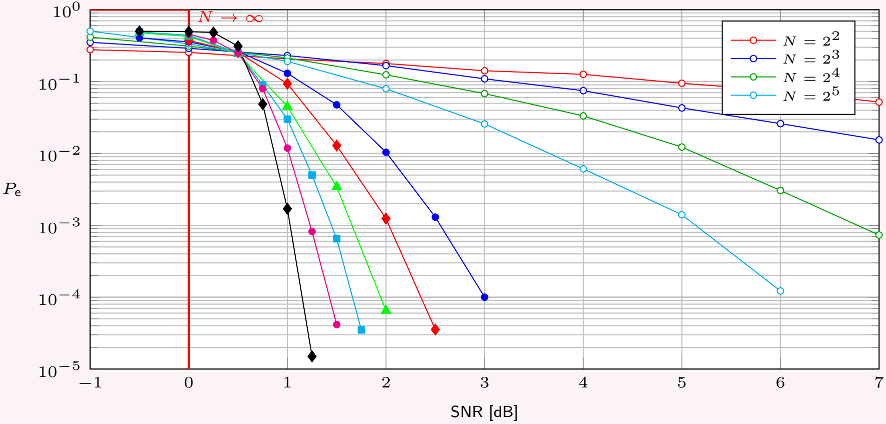
[1] C. Berrou et al., “Near Shannon limit error-correcting coding and decoding: Turbo-codes,” in Proc. IEEE ICC 1993.

[2] E. Arikan, “Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels,” in IEEE Trans. Inf. Theory, July 2009.

Example: Polar Codes, $R_N = 1/2$

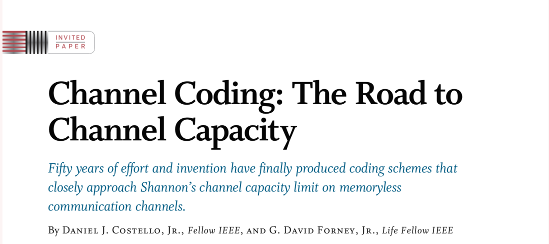


Example: Random Codes vs. Polar Codes



Channel Coding: The Road to Channel Capacity

Turbo Codes vs. NASA Codes

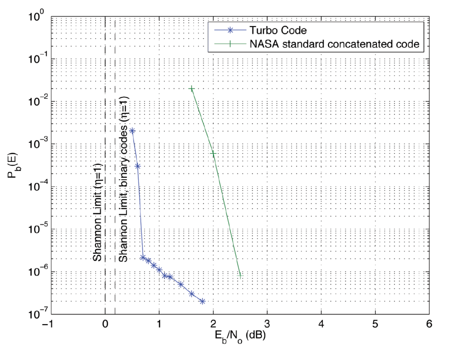


INVITED PAPER

Channel Coding: The Road to Channel Capacity

Fifty years of effort and invention have finally produced coding schemes that closely approach Shannon's channel capacity limit on memoryless communication channels.

By Daniel J. Costello, Jr., *Fellow IEEE*, and G. David Forney, Jr., *Life Fellow IEEE*



Excellent review paper [3] where coding history is explained, including NASA codes and the so-called turbo revolution

All this and more in 5LSF0, "Applications of Information Theory" (Q4) and "Information Theory" 5XSE0 (Q3)

[3] Daniel J. Costello, Jr., and G. D. Forney, Jr., "Channel Coding: The Road to Channel Capacity," Proc. of the IEEE, vol. 95, no. 6, June 2007.

Module 9.2

Presentation Outline

Capacity Proof

Sketch of the Proof

Four key ingredients:

Part 1: Sphere Hardening of Gaussian Vectors

Part 2: Random Coding Generation

Part 3: Everything is Gaussian

Part 4: Error Probability

Part 1: Sphere Hardening of Gaussian Vectors (1/2)

Gaussian Vectors

Consider a random Gaussian vector $\underline{G}$ with N components, each with mean 0 and variance σ_g^2 and a normalized version of this vector: $\underline{G}' = \underline{G}/\sqrt{N}$. The square norms (lengths) of these vectors ($\|\underline{G}\|^2$ and $\|\underline{G}'\|^2$) are random variables.

Sphere Hardening of Gaussian Vectors

The normalized vector $\underline{G}'$ can be shown to have the following mean and variance:

$$E\left[\|\underline{G}'\|^2\right] = \sigma_g^2, \qquad \text{var}\left[\|\underline{G}'\|^2\right] = \frac{2\sigma_g^4}{N}.$$

Thus, vectors $\underline{G}'$ are **on the surface of a hypersphere** with radius σ_g. Fluctuations are possible, however for $N \to \infty$ these fluctuations disappear.

Part 1: Sphere Hardening of Gaussian Vectors (2/2)

Part 2: Random Code Generation

Generating a Random Code

- Fix the number of signal vectors $|\mathcal{M}|$ and their number of components (dimensions) N
- Select $|\mathcal{M}|$ signal vectors $\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{|\mathcal{M}|}$ *at random*, independently of each other
- Make each vector component is a random sample from a Gaussian density with mean 0 and variance E_N:
  - By the sphere-hardening argument, *energies of the vectors* $E[\|\underline{S}\|^2]$ are actually roughly NE_N.
  - The *expected energy per dimension* is E_N.
- Consider the *ensemble* of all signal sets that can be chosen in this way

Part 3: Everything is Gaussian

Gaussian Inputs, Gaussian Outputs

- The channel output vector is $\underline{r} = \underline{s}_m + \underline{n}$
- The components of the noise vector $\underline{n}$ are Gaussian with mean 0 and variance $N_0/2$
- The sum of two independent Gaussian vectors is also Gaussian, with all components having mean 0 and variance $E_N + N_0/2$

Consider Normalized Quantities with $N \to \infty$

- Output is $\underline{r}' = \underline{s}'_m + \underline{n}'$, where $\underline{s}'_m = \underline{s}_m/\sqrt{N}$ and $\underline{n}' = \underline{n}/\sqrt{N}$
- Normalized vector $\underline{s}'_m$ is on the surface of a hypersphere with radius $\sqrt{E_N}$, i.e., $\|\underline{s}'_m\|^2 = E_N$
- Normalized noise vector $\underline{n}'$ is on the surface of a hypersphere with radius $\sqrt{N_0/2}$, i.e., $\|\underline{n}'\|^2 = N_0/2$
- Normalized received vector $\underline{r}'$ is on the surface of a hypersphere with radius $\sqrt{E_N + N_0/2}$, i.e., $\|\underline{r}'\|^2 = E_N + N_0/2$

Part 4: Error Probability

Average Error Probability

- We are interested in $P_{\mathsf{e}}^{\mathsf{av}}$, i.e., the error probability P_{e} *averaged over the ensemble of signal sets*, i.e.,

$$P_{\mathsf{e}}^{\mathsf{av}} = \int_{\mathbb{R}^{N\cdot|\mathcal{M}|}} p(\underline{s}_1, \underline{s}_2, \dots, \underline{s}_{|\mathcal{M}|}) P_{\mathsf{e}}(\underline{s}_1, \underline{s}_2, \dots, \underline{s}_{|\mathcal{M}|}) d\underline{s}_1 d\underline{s}_2 \dots d\underline{s}_{|\mathcal{M}|}$$

- Once we know $P_{\mathsf{e}}^{\mathsf{av}}$ we claim that there exists *at least one signal set* $\{\underline{s}_1, \underline{s}_2, \dots, \underline{s}_{|\mathcal{M}|}\}$ with error probability $P_{\mathsf{e}}(\underline{s}_1, \underline{s}_2, \dots, \underline{s}_{|\mathcal{M}|}) \leq P_{\mathsf{e}}^{\mathsf{av}}$
- It can then be shown that if $R_N = C_N - \delta$

$$\lim_{N\to\infty} P_{\mathsf{e}}^{\mathsf{av}} \leq \lim_{N\to\infty} 2^{-\delta N} \sqrt{\frac{E_N + N_0/2}{N_0/2}} = 0$$

Summary Module 9.2

Take Home Messages

- Four key ingredients in the capacity proof
- Codes with Gaussian inputs achieve capacity
- Modern codes exist that approach capacity
- Is channel capacity mportant? Yes because...
  - It is a fundamental limit with beautiful and simple derivations
  - We want to approach such limits
  - Signal shaping

Communication Theory (5ETB0)
Module 10.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 10.1

Presentation Outline

Problem Description: Serial Transmission

Motivation

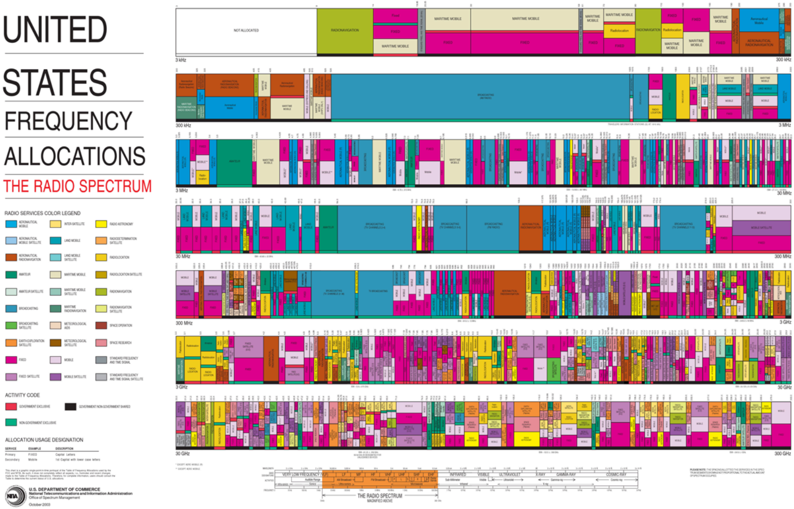
- A channel with a bandwidth of W Hz can accommodate roughly $2WT$ dimensions each T seconds (dimensionality theorem)
- Obtained for building-block waveforms that are zero outside $[0, T]$
- Can we get $2WT$ extra dimensions every new T seconds ($2W$ extra dimensions per second)?
  - Answer is yes, using building-blocks waveforms with nonfinite duration

Consequences

- Building blocks that are not time-limited:
  - Finite bandwidth, but
  - inter-symbol interference is created

In this module we will show that time-shifted versions of the original pulse can be used, and that some properties of the pulse make the interference disappear (with the right receiver).

Motivation: Who Cares about the Spectrum?



Module 10.1

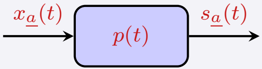
Presentation Outline

A Question Before Getting Started...

A linear filter fed with a train of impulses

- Suppose we have a linear filter
- Impulse response $p(t)$ and fed with a train of weighted impulses

$$x_{\underline{a}}(t) = \sum_{k=0}^{K-1} a_k \delta(t - kT)$$



where $\underline{a} = (a_0, a_1, \dots, a_{K-1})$

- What is the output of the filter $s_{\underline{a}}(t)$?

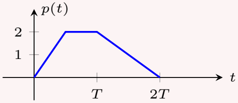
Answer:

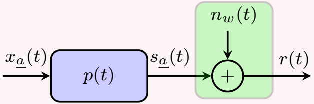
$$s_{\underline{a}}(t) = \sum_{k=0}^{K-1} a_k p(t - kT)$$

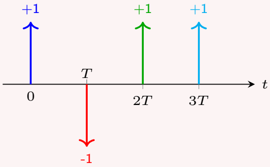
Model: Binary PAM

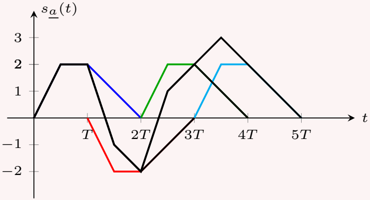
Serial pulse-amplitude modulation (PAM)

Transmitted signal is: $s_{\underline{a}}(t) = \sum_{k=0}^{K-1} a_k p(t - kT)$. Vector of amplitudes $\underline{a} = (a_0, a_1, \ldots, a_{K-1})$ consists of symbols $a_k, k = 0, \ldots, K-1$ taking values in the alphabet $\mathcal{A} = \{-1, +1\}$





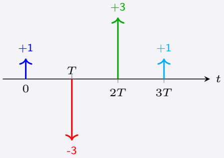


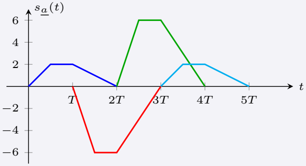


Model: Nonbinary PAM

Non-binary PAM Example

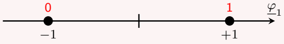
To double the transmission rate, we can double the number of bits per symbol. This can be obtained using an alphabet with 4 messages: $\mathcal{A} = \{-3, -1, 1, 3\}$.

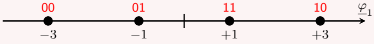




Two Comments

- Serial PAM is similar to bit-by-bit signaling with $K = 1$. Differences: not using time-limited pulses and nonbinary alphabets $\mathcal{A}$.
- Geometrical representations of binary and nonbinary PAM are:





Summary Module 10.1

Take Home Messages

- Pulses that are not time-limited can be used for transmission
- Serial pulse amplitude modulation (binary and nonbinary)
- Inter-symbol interference is generated

Communication Theory (5ETB0)
Module 10.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

Module 10.2

Presentation Outline

Orthonormal Pulses: the Nyquist Criterion

Problem and Solution

Problem: Inter-symbol interference.
Solution: Pulses $p(t)$ such that all time shifts by T [s] of the $p(t)$ form an orthonormal basis.

The Nyquist Result in the Time domain

- Pulse $p(t)$ has to satisfy for integer k and k'

$$\int_{-\infty}^{\infty} p(t-kT)p(t-k'T)dt = \begin{cases} 1 & \text{if } k = k' \\ 0 & \text{if } k \neq k' \end{cases}$$

- Equivalently:

$$\int_{-\infty}^{\infty} p(\alpha)p(\alpha-kT)d\alpha = p(t) * p(-t)\Big|_{t=kT} = h(kT) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$$

  where $h(t) \triangleq p(t) * p(-t)$.
- Time-domain restriction on the pulse $p(t)$ is called zero-forcing (ZF) criterion

Orthonormal Pulses: the Nyquist Criterion

The Nyquist Result in the Frequency Domain

Let $H(f)$ be the Fourier transform of $h(t) = p(t) * p(-t)$, where $H(f) = P(f)P^*(f) = |P(f)|^2$. The Nyquist criterion in the frequency domain is then

$$Z(f) = \frac{1}{T} \sum_{m=-\infty}^{\infty} H(f + m/T) = 1 \qquad \text{for all } f,$$

or equivalently

$$= \frac{1}{T} \sum_{m=-\infty}^{\infty} |P(f + m/T)|^2 = 1 \qquad \text{for all } f.$$

Two Comments

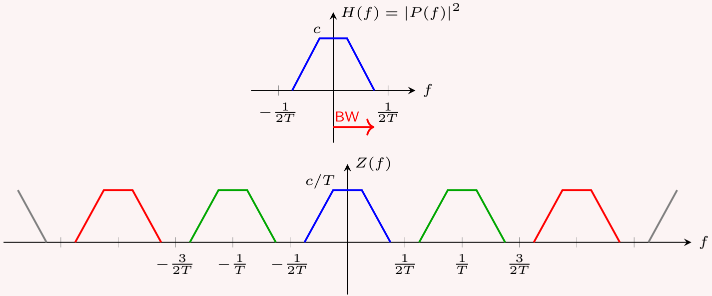
- Note that the Fourier transform of $p(t) * p(-t)$ is $|P(f)|^2$ because $p(t)$ is real
- When discussing orthogonality, T is important (T-orthogonality)

Module 10.2

Presentation Outline

Case 1: Negative

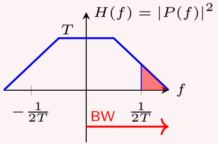
Negative Case: $W < 1/(2T)$

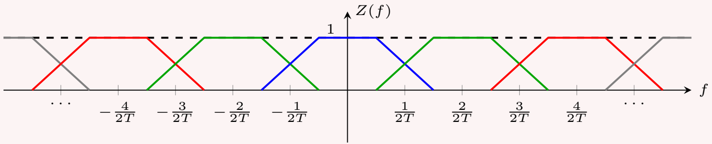


- Bandwidth W of the pulse $p(t)$ is strictly smaller than $1/(2T)$
- Nyquist criterion $(H(f) = |P(f)|^2)$ can not be satisfied

Case 2: Excess

Excess Bandwidth Case $W > 1/(2T)$





- Bandwidth W of the pulse $p(t)$ is larger than $1/(2T)$
- *Excess bandwidth*: bandwidth minus $1/(2T)$
- Popular pulse with an excess bandwidth (yet orthonormal): square root-raised cosine

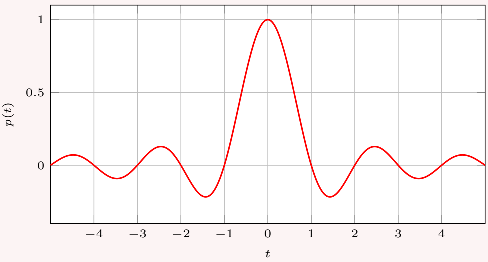
Case 3: Match (1/2)

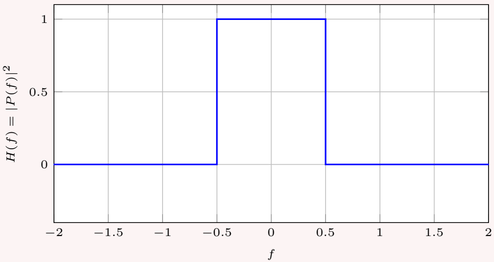
Matching Case $W = 1/(2T)$ (1/2)

- Basic pulse $p(t)$ that corresponds to the ideally bandlimited spectrum is the sinc-pulse:

$$p(t) = \frac{1}{\sqrt{T}} \frac{\sin(\pi t/T)}{\pi t/T}$$

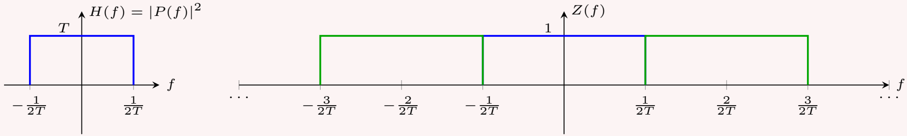
- For $T = 1$





Case 3: Match (2/2)

Matching Case $W = 1/(2T)$ (2/2)



- Smallest possible bandwidth W of a pulse that satisfies the Nyquist criterion is $1/(2T)$
- "Basic" pulse $P(f)$ with $W = 1/(2T)$ for which the Nyquist criterion holds has a brick-wall (ideally bandlimited) spectrum
- Basic pulse $p(t)$ that corresponds to the ideally bandlimited spectrum is the sinc-pulse

$$P(f) = \begin{cases} \sqrt{T} & \text{if } |f| < 1/(2T), \\ 0 & \text{if } |f| > 1/(2T) \end{cases}$$

Pulses with smallest Bandwidth

Sinc Pulses

The smallest possible bandwidth W of a pulse that satisfies the Nyquist criterion is $W = \frac{1}{2T}$. The sinc-pulse $p(t) = \frac{1}{\sqrt{T}} \frac{\sin(\pi t/T)}{\pi t/T}$ has this property. This way of serial pulse transmission leads to exactly $\frac{1}{T} = 2W$ dimensions per second.

Extra Considerations

Q1: What problems do sinc-pulses cause in practice?

- Infinitely long
- Difficult to generate
- Noncausal

Q2: What are the positive aspects of sinc pulses?

- Orthogonal pulses (no ISI)
- Best possible use of bandwidth
- Only one building-block waveform (one new dimension) every T [s]

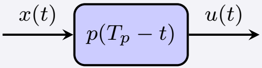
Module 10.2

Presentation Outline

Before Getting Started...

A Matched Filter

- Suppose we have a linear filter with impulse response $h(t) = p(T_p - t)$ and fed with a signal $x(t)$.
- Delay T_p is to make the filter causal



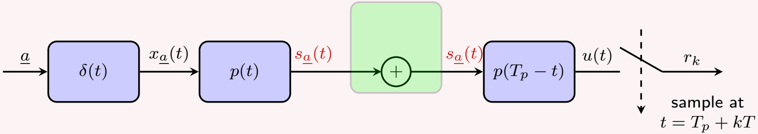
Q1: What is the output of the filter $u(t)$?

Answer:

$$u(t) = \int_{-\infty}^{\infty} x(\tau)p(\tau - t + T_p)d\tau$$

Optimum Receiver: Without Noise

Serial PAM Transceiver



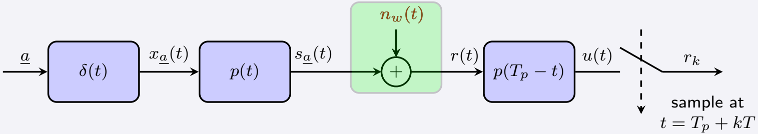
$$r_k = u(T_p + kT) = \int_{-\infty}^{\infty} r(\tau)p(T_p - t + \tau)d\tau\Big|_{t=T_p+kT} = \int_{-\infty}^{\infty} r(\tau)p(-kT + \tau)d\tau$$

$$= \int_{-\infty}^{\infty} \left(\sum_{k'=0}^{K-1} a_{k'} p(\tau - k'T)\right) p(-kT + \tau)d\tau$$

$$= \sum_{k'=0}^{K-1} a_{k'} \underbrace{\int_{-\infty}^{\infty} p(\tau - k'T)p(-kT + \tau)d\tau}_{0 \text{ unless } k' = k} = a_k$$

Optimum Receiver: With Noise

Serial PAM Transceiver



$$r_k = u(T_p + kT) = \int_{-\infty}^{\infty} r(\tau) p(T_p - t + \tau) d\tau \Big|_{t = T_p + kT} = \int_{-\infty}^{\infty} r(\tau) p(-kT + \tau) d\tau$$

$$= \int_{-\infty}^{\infty} \left(\sum_{k'=0}^{K-1} a_{k'} p(\tau - k'T) + n_w(\tau) \right) p(-kT + \tau) d\tau$$

$$= a_k + \int_{-\infty}^{\infty} n_w(\tau) p(-kT + \tau) d\tau = a_k + z_k \qquad z_k \sim \mathcal{N}(0, N_0/2)$$

Serial PAM Transceiver

Concluding Remarks

- (Serial) PAM transceiver uses a single filter (very simple!)
- No ISI is present at the receiver
- If $p(t)$ are sinc pulses: best use of bandwidth
- Abstract waveform transmission into DICO channel: optimum detection

Summary Module 10.2

Take Home Messages

- PAM transceiver
- Nyquist criterion (time and frequency): zero ISI!
- Bandwidth considerations: Negative, excess and matching

Communication Theory (5ETB0)
Module 11.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

Module 11.1

Presentation Outline

Part I Motivation

Part II Passband and Quadrature Multiplexing

Part III Capacity of the Passband Channel

Passband Transmission: Motivation (1/2)

Baseband Transmission

- For a baseband pulse $p(t)$ with Fourier transform $P(f)$, the BW of the transmitted signal is at least $1/2T$, where $1/T$ is the symbol rate
- Best case scenario, we use pulses sinc pulses and the BW is exactly $1/2T$
- Baseband signals can be sent over a telephone line or a coaxial cable

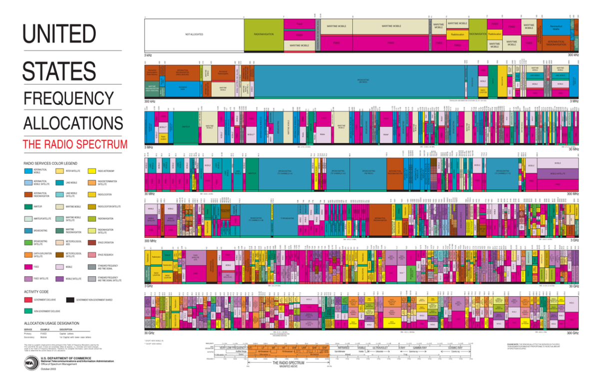
Passband Transmission

- FM radio operates at 100 [MHz], $\lambda = v/f \approx 3$ m $\Longrightarrow$ Large antennas
- WiFi operates at 2.4 [GHz] or 5 [GHz], $6 \lessapprox \lambda \lessapprox 12.5$ cm $\Longrightarrow$ Small antennas
- More transmission bandwidth available at higher frequencies.

Passband Transmission: Motivation (2/2)

Other Reasons for Passband Transmission

- Makes best use of the channel
- Allows us to assign different users to different frequencies



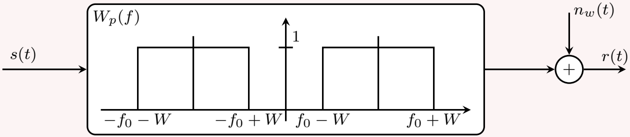
Module 11.1

Presentation Outline

System Model

Ideal passband channel with AWGN

- Ideal passband filter



$$W_p(f) \triangleq \begin{cases} 1 & \text{if } f_0 - W < |f| < f_0 + W \\ 0 & \text{elsewhere} \end{cases}$$

- Noise is AWGN, PSD is:

$$S_{N_w}(f) = N_0/2 \text{ for } -\infty < f < \infty$$

Quadrature Multiplexing: Model

Quadrature Multiplexing Transmitter

- Consider two set of baseband waveforms having bandwidth smaller than W:

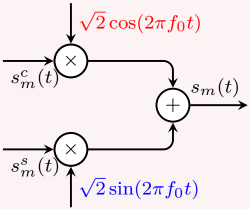
$$\{s_1^c(t), s_2^c(t), \ldots, s_{|\mathcal{M}|}^c(t)\}$$

$$\{s_1^s(t), s_2^s(t), \ldots, s_{|\mathcal{M}|}^s(t)\}$$

- Corresponding building-block waveforms are:

$$\phi_i(t), i = 1, 2, \ldots, N_c, \quad N_c \leq |\mathcal{M}|$$

$$\psi_j(t), j = 1, 2, \ldots, N_s, \quad N_s \leq |\mathcal{M}|$$



Transmitted signal is:

$$s_m(t) = s_m^c(t)\sqrt{2}\cos(2\pi f_0 t) + s_m^s(t)\sqrt{2}\sin(2\pi f_0 t)$$

Quadrature Multiplexing: Orthogonality (1/2)

Transmitted waveform

Building-block waveforms $\phi_i(t)$ and $\psi_j(t)$:

$$s^c_m(t) = \sum_{i=1}^{N_c} s^c_{mi}\phi_i(t) \quad \text{and} \quad s^s_m(t) = \sum_{j=1}^{N_s} s^s_{mj}\psi_j(t)$$

The mth transmitted waveform is

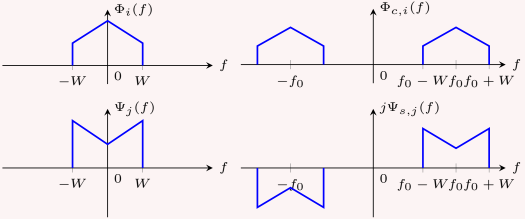
$$\begin{aligned} s_m(t) &= s^c_m(t)\sqrt{2}\cos(2\pi f_0 t) + s^s_m(t)\sqrt{2}\sin(2\pi f_0 t) \\ &= \sum_{i=1}^{N_c} s^c_{mi}\phi_i(t)\sqrt{2}\cos(2\pi f_0 t) + \sum_{j=1}^{N_s} s^s_{mj}\psi_j(t)\sqrt{2}\sin(2\pi f_0 t) \\ &= \sum_{i=1}^{N_c} s^c_{mi}\phi_{c,i}(t) + \sum_{j=1}^{N_s} s^s_{mj}\psi_{s,j}(t) \end{aligned}$$

where $\phi_{c,i}(t)$ and $\psi_{s,j}(t)$ are in-phase and quadrature building-block waveforms.

Quadrature Multiplexing: Orthogonality (2/2)

Building-block waveforms?

Are $\phi_{c,i}(t)$ and $\psi_{s,j}(t)$ building-block waveforms?
Step 1 of proof is (11.9) of the reader:



Step 2 of proof is to show (see (11.10)–(11.12)) of the reader)

$$\int_{-\infty}^{\infty} \phi_{c,i}(t)\phi_{c,i'}(t)dt = \int_{-\infty}^{\infty} \Phi_{c,i}(f)\Phi^*_{c,i'}(f)df = \delta_{i,i'}$$

$$\int_{-\infty}^{\infty} \phi_{s,i}(t)\phi_{s,i'}(t)dt = \delta_{i,i'} \quad \text{and} \quad \int_{-\infty}^{\infty} \phi_{s,i}(t)\phi_{c,i'}(t)dt = 0$$

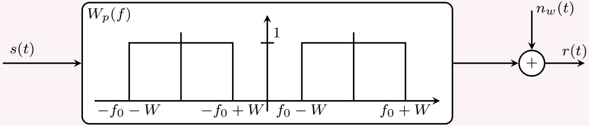
Quadrature Multiplexing: Building-block Waveforms

Quadrature Building-Block Waveforms

We have shown that all in-phase building-block waveforms $\phi_{c,i}(t)$ for $i = 1, ..., N_c$ and all quadrature building-block waveforms $\psi_{s,j}(t)$ for $j = 1, ..., N_s$ together form an orthonormal basis.

Bandwidth Considerations

The spectra $\Phi_{c,i}(f)$ and $\Psi_{s,j}(f)$ of all these building-block waveforms are zero outside the passband $\pm[f_0 - W, f_0 + W]$. Therefore none of these building-block waveforms is hindered by the passband filter $W(f)$ when they are sent over our passband channel.



Module 11.1

Presentation Outline

Capacity of the Passband Channel

Passband Channel Capacity

The capacity is $C_N = 1/2 \log_2 (1 + E_N/(N_0/2))$ [bit/dimension], where $E_N = P_s/(4W)$ [Joule/dimension] and $4W$ is the number of dimension/second for the passband channel (dimensionality theorem).
The capacity of the passband channel with bandwidth $\pm[f_0 - W, f_0 + W]$ is

$$C_N = \frac{1}{2} \log_2 \left(1 + \frac{P_s}{2N_0 W}\right) \quad \left[\frac{\text{bit}}{\text{dimension}}\right]$$

The capacity in bit per second is

$$C = 2W \log_2 \left(1 + \frac{P_s}{2N_0 W}\right) \quad \left[\frac{\text{bit}}{\text{second}}\right]$$

Connection to Baseband AWGN Capacity

The passband bandwidth is $2W$, thus

$$C = W \log_2 \left(1 + \frac{P_s}{W N_0}\right) \quad \left[\frac{\text{bit}}{\text{second}}\right].$$

Who Cares?

IEEE 802.11 Standards (WiFi)

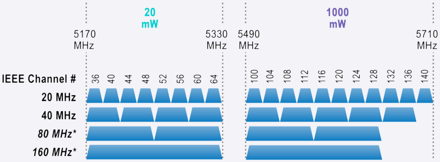


FIGURE 2 • *European channel allocations for 20/40/80/160 MHz*

Summary Module 11.1

Take Home Messages

- Motivation for passband transmission
- Quadrature transmitter using in-phase and quadrature building-blocks
- Capacity of the passband channel has the same structure of the baseband AWGN channel

Communication Theory (5ETB0)
Module 11.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands
www.tue.nl/ictlab/

Module 11.2

Presentation Outline

Quadrature Multiplexing: Optimum Receiver (1/3)

Passband Transmitted Waveform

$$s_m(t) = \sum_{i=1}^{N_c} s^c_{mi}\phi_{c,i}(t) + \sum_{j=1}^{N_s} s^s_{mj}\psi_{s,j}(t)$$

Optimum Receiver

The optimum receiver applies the rule

$$\hat{m}^{\text{MAP}}(\underline{r}) = \operatorname*{argmax}_{m\in\mathcal{M}}\{(\underline{r}\cdot\underline{s}_m) + c_m\}$$

where

$$c_m = \frac{N_0}{2}\ln\Pr\{M=m\} - \frac{E_m}{2}$$

and E_m is the energy of the waveform $s_m(t)$, for $m \in \mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$.

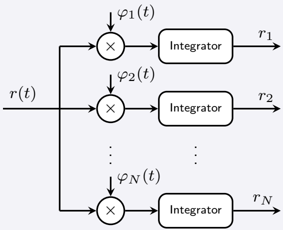
Recap: Correlation Receiver

Correlation Receiver

The transmitted waveform is

$$s_m(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t)$$

where $\varphi_i(t)$ are N building-block waveforms.
The received waveform is $r(t) = s_m(t) + n_w(t)$.



Q1: What is the structure of the corresponding correlation receiver that gives us the r-values $\underline{r} = (r_1, r_2, \ldots, r_N)$?
Answer: N multipliers and integrators

Quadrature Multiplexing: Optimum Receiver (2/3)

Computing the r-values

$$\begin{aligned}\int_{-\infty}^{\infty} r(t)\phi_{c,i}(t)dt &= \int_{-\infty}^{\infty} r(t)\sqrt{2}\cos(2\pi f_0 t)\phi_i(t)dt = r_i^c, \quad i = 1,2,\ldots,N_c\\ \int_{-\infty}^{\infty} r(t)\psi_{s,j}(t)dt &= \int_{-\infty}^{\infty} r(t)\sqrt{2}\sin(2\pi f_0 t)\psi_j(t)dt = r_j^s, \quad i = 1,2,\ldots,N_s\end{aligned}$$

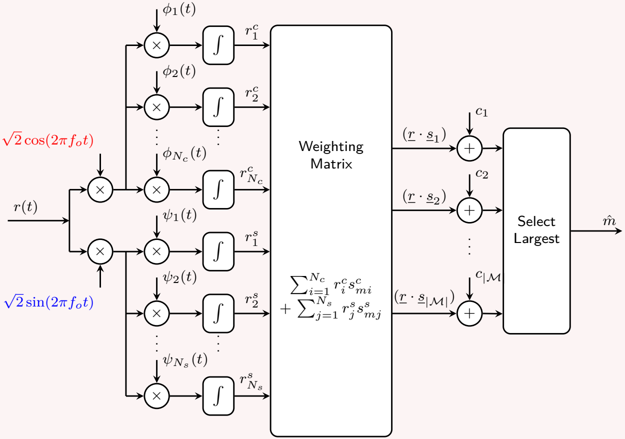
Computing the dot products and constants

$$(\underline{r}\cdot\underline{s}_m) = \sum_{i=1}^{N_c} r_i^c s_{mi}^c + \sum_{j=1}^{N_s} r_j^s s_{mj}^s$$

$$c_m = \frac{N_0}{2}\ln \Pr\{M=m\} - \frac{\|\underline{s}_m\|^2}{2}, \qquad \|\underline{s}_m\|^2 = \|\underline{s}_m^c\|^2 + \|\underline{s}_m^s\|^2$$

Quadrature Multiplexing: Optimum Receiver (3/3)

Computing the r-values



Quadrature Multiplexing: Short Pause

What have we done so far?

The mth transmitted waveform is

$$s_m(t) = \sum_{i=1}^{N_c} s_{mi}^c \phi_i(t)\sqrt{2}\cos(2\pi f_0 t) + \sum_{j=1}^{N_s} s_{mj}^s \psi_j(t)\sqrt{2}\sin(2\pi f_0 t)$$

- Symbols $\underline{s}_m^c$ and $\underline{s}_m^s$ modulated using $\phi_i(t)$ and $\psi_j(t)$
- The resulting signal is BW-limited (to W)
- Then up-conversion around frequency f_0
- At Rx: down-conversion, followed by correlation receiver

Questions:

Q1: What is the dimensionality of the signal space in the general case above?
Q2: How do we make this complex system to be baseband PAM?
Q3: What is the dimensionality of the signal space now?

Module 11.2

Presentation Outline

Quadrature Amplitude Modulation

Simplifying general model gives us Quadrature Amplitude Modulation (QAM)

The mth transmitted waveform is

$$s_m(t) = \sum_{i=1}^{N_c} s_{mi}^c \phi_i(t)\sqrt{2}\cos(2\pi f_0 t) + \sum_{j=1}^{N_s} s_{mj}^s \psi_j(t)\sqrt{2}\sin(2\pi f_0 t)$$

- Make $N_c = N_s = N$ and the baseband in-phase and quadrature building-block waveforms to be equal ($\phi_i(t) = \psi_i(t)$ for all $i = 1, \ldots, N$):

$$s_m(t) = \sum_{i=1}^{N} \phi_i(t)\left(s_{mi}^c \sqrt{2}\cos(2\pi f_0 t) + s_{mi}^s \sqrt{2}\sin(2\pi f_0 t)\right)$$

- Take a single dimension ($N = 1$) with a and b amplitudes for in-phase and quadrature:

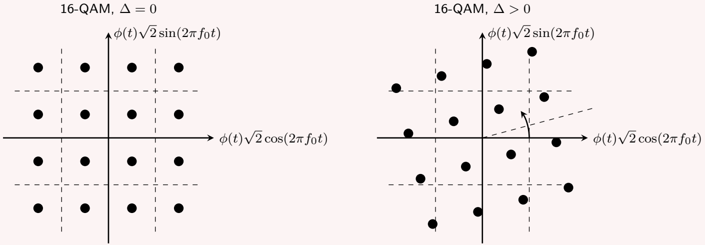
$$s_m(t) = a\phi(t)\sqrt{2}\cos(2\pi f_0 t) + b\phi(t)\sqrt{2}\sin(2\pi f_0 t)$$

Quadrature Amplitude Modulation

Time Delay $\Rightarrow$ Phase Rotation

What happens if the channel introduces delay? If we observe a slightly delayed version of the signal $r(t) = s(t-\Delta) + n_w(t)$:

$$
\begin{aligned}
s(t-\Delta) \quad \approx \quad & [a\cos(2\pi f_0 \Delta) - b\sin(2\pi f_0\Delta)]\, \phi(t)\sqrt{2}\cos(2\pi f_0 t) \\
& + [b\cos(2\pi f_0 \Delta) + a\sin(2\pi f_0\Delta)]\, \phi(t)\sqrt{2}\sin(2\pi f_0 t).
\end{aligned}
$$



Module 11.2

Presentation Outline

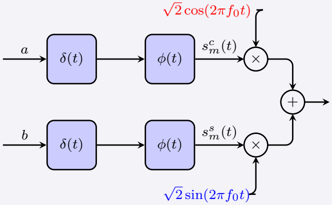
Part I Quadrature Multiplexing Receiver

Part II Quadrature amplitude modulation (QAM)

Part III Serial QAM

From QAM to Serial QAM

Quadrature Amplitude Multiplexing so far

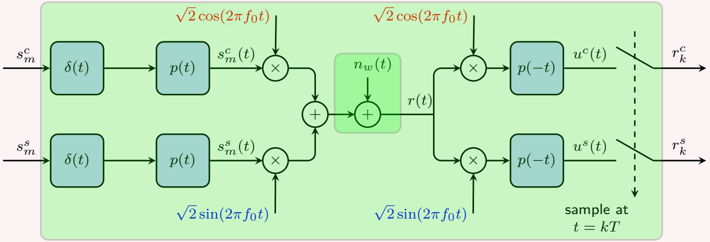


Serial QAM

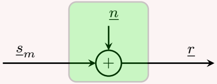
- Use serial PAM with a pulse $\phi(t) = p(t)$ satisfying the Nyquist criterion
- Use same pulse in cosine and sine branches
- Use sinc pulses as baseband building blocks with T [s] shifts

Serial QAM Transceiver

Serial QAM System



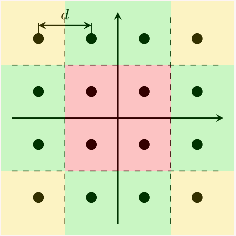
The whole transmission chain can be replaced by a Vector (2D) DICO channel



$$\underline{r} = \underline{s}_m + \underline{n}$$
$$\underline{s}_m = (s_m^c, s_m^s)$$
$$n_i \sim \mathcal{N}(0, N_0/2), i = 1, 2$$

Error Probability for 16-QAM (1)

Serial QAM Error Probability



Union bound

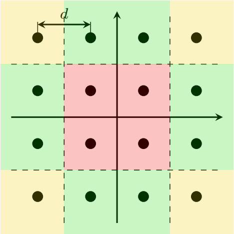
- Yellow points: $P_\mathrm{e}^{Y} \leq 2Q\left(\frac{d}{2\sigma}\right)$
- Green points: $P_\mathrm{e}^{G} \leq 3Q\left(\frac{d}{2\sigma}\right)$
- Red points: $P_\mathrm{e}^{R} \leq 4Q\left(\frac{d}{2\sigma}\right)$

Exact error probability for each point

- Yellow points: $P_e^{Y} = 2Q\left(\frac{d}{2\sigma}\right) - Q^2\left(\frac{d}{2\sigma}\right)$
- Green points: $P_e^{G} = 3Q\left(\frac{d}{2\sigma}\right) - 2Q^2\left(\frac{d}{2\sigma}\right)$
- Red points: $P_e^{R} = 4Q\left(\frac{d}{2\sigma}\right) - 4Q^2\left(\frac{d}{2\sigma}\right)$

Error Probability for 16-QAM (2)

Serial QAM Error Probability



Upper bound on the total average error probability:

$$
\begin{aligned}
P_{\mathrm{e}} \leq & \frac{4}{16} \cdot 2Q\left(\frac{d}{2\sigma}\right) + \frac{8}{16} \cdot 3Q\left(\frac{d}{2\sigma}\right) + \\
& \frac{4}{16} \cdot 4Q\left(\frac{d}{2\sigma}\right) \\
= & \frac{48}{16} Q\left(\frac{d}{2\sigma}\right)
\end{aligned}
$$

Who Cares?

Modulation Enhancements

Like most recent wireless specification, 802.11ac uses Orthogonal Frequency-Division Multiplexing (OFDM) to modulate bits for transmission over the wireless medium. While the modulation approach is identical to that used in 802.11n, 802.11ac optionally allows the use of 256 QAM in addition to the mandatory Quadrature Phase Shift Keying (QPSK), Binary PSK (BPSK), 16 QAM and 64 QAM modulations. 256 QAM increases the number of bits per sub-carrier from 6 to 8, resulting in a 33% increase in PHY rate under the right conditions. It should be noted however that 256 QAM can only be used in high signal-to-noise ratio (SNR) scenarios (across the used spectrum and desired streams); i.e. for very favorable channel conditions. The support of 256 QAM will increase the maximum PHY rate that can be supported by the system, but will have no effect in typical scenarios and will not lead to any reach increase for the service. Also, supporting 256 QAM requires transmitter and receiver to be designed such that the inherent SNR (transmit and receive

Summary Module 11.2

Take Home Messages

- Receiver structure for quadrature multiplexing
- Quadrature amplitude modulation
- Serial QAM System: Continuous-time vs. discrete-time

Communication Theory (5ETB0)
Module 12.1

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

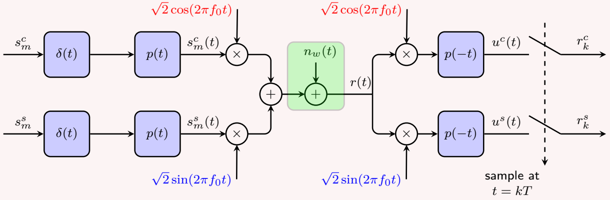
Module 12.1

Presentation Outline

Motivation: Coherent vs. Incoherent Communications

Coherent Reception

- *Coherent* reception: Phase of Tx and Rx are equal
- Very popular case: Serial QAM



Potential Problems and Solution

- Frequency mismatch: Diffent Tx an Rx oscilators
- Phase mismatch can be caused by propagation delay

Incoherent reception: Phase difference is unknown and/or cannot be compensated

Model

System Model

- Transmitted signal $s(t) = s^b(t)\sqrt{2}\cos(2\pi f_0 t - \theta)$
- Signal vectors are $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$, where

$$s_m^b(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t).$$

- Phase θ is assumed random and uniform over $[0, 2\pi)$:

$$p_\Theta(\theta) = \frac{1}{2\pi}, \text{ for } 0 \leq \theta < 2\pi.$$

A Few Assumptions...

- Phase difference assumed at the transmitter (without loss of generality)
- Spectrum of baseband signal within $[-W, W]$
- Only one baseband signal (for simplicity)

Quadrature-multiplexed Transmitted Waveform

Connection to QAM: N or $2N$ Dimensions?

$$
\begin{aligned}
s_m(t) &= s_m^b(t)\sqrt{2}\cos(2\pi f_0 t - \theta)\\
&= s_m^b(t)\cos(\theta)\sqrt{2}\cos(2\pi f_0 t) + s_m^b(t)\sin(\theta)\sqrt{2}\sin(2\pi f_0 t)\\
&= \sum_{i=1}^{N} s_{mi}^c \varphi_i(t)\sqrt{2}\cos(2\pi f_0 t) + \sum_{i=1}^{N} s_{mi}^s \varphi_i(t)\sqrt{2}\sin(2\pi f_0 t)\\
&= \sum_{i=1}^{N} s_{mi}^c \phi_{c,i}(t) + \sum_{i=1}^{N} s_{mi}^s \psi_{s,i}(t)
\end{aligned}
$$

where

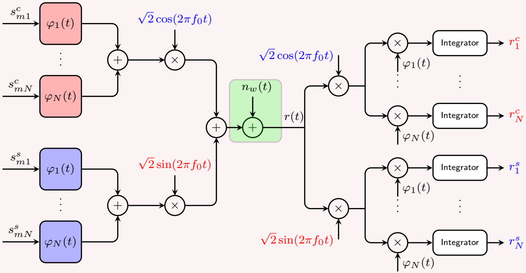
$$
\begin{aligned}
\underline{s}_m^c &= (s_{m1}, s_{m2}, \ldots, s_{mN})\cos(\theta) = \underline{s}_m \cos(\theta)\\
\underline{s}_m^s &= (s_{m1}, s_{m2}, \ldots, s_{mN})\sin(\theta) = \underline{s}_m \sin(\theta)
\end{aligned}
$$

Module 12.1

Presentation Outline

Optimum Incoherent Reception (1/3)

Forming the Vectors



Forming the Vectors

An optimum-receiver forms $\underline{r}^c$ and $\underline{r}^s$ using $r(t) = s_m(t) + n_w(t)$:

$$\begin{aligned} \underline{r}^c &= \underline{s}_m^c + \underline{n}^c = \underline{s}_m \cos\theta + \underline{n}^c \\ \underline{r}^s &= \underline{s}_m^s + \underline{n}^s = \underline{s}_m \sin\theta + \underline{n}^s \end{aligned}$$

Optimum Incoherent Reception (2/3)

Optimum Receiver for Equally Likely Messages (Result 12.1)

The optimum receiver for "random-phase transmission" (incoherent detection) is:

$$\hat{m}^{\mathsf{MAP}}(\underline{r}) = \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{ I_0\left(\frac{2X_m}{N_0}\right) \exp\left(-\frac{E_m}{N_0}\right) \right\}$$

where

$$X_m \triangleq \sqrt{(\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2}$$

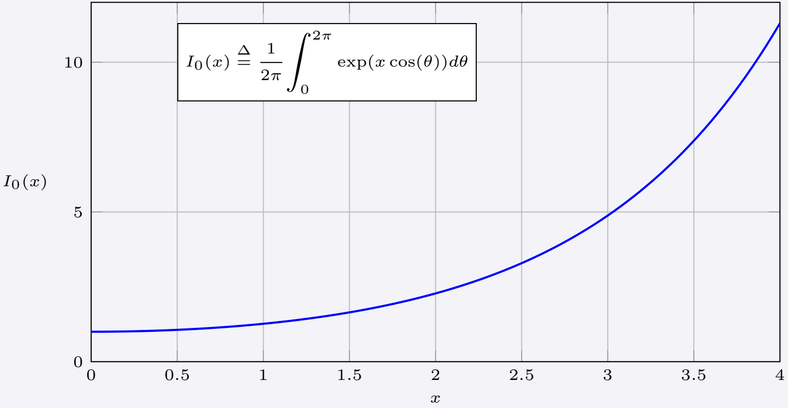
and E_m is the energy of the m-th message

Proof Sketch

- Compute MAP decision variables for a given θ, average over the distribution of Θ
- Simplify expressions (details in Sec. 12.2 for details)
- Express result using the zero-order modified Bessel function of the first kind

Zero-order modified Bessel function of the first kind

$I_0(x)$ is an increasing function of x

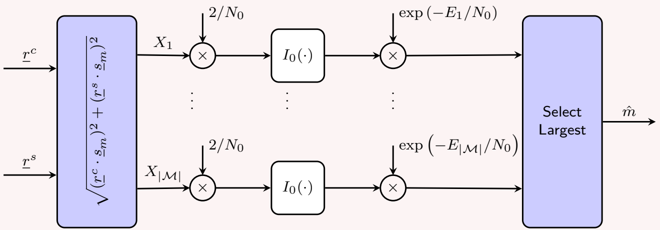


Optimum Incoherent Reception (3/3)

Optimum Receiver for Equally Likely Messages

$$I_0\left(\frac{2X_m}{N_0}\right)\exp\left(-\frac{E_m}{N_0}\right), \quad X_m \triangleq \sqrt{(\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2}$$

General Structure

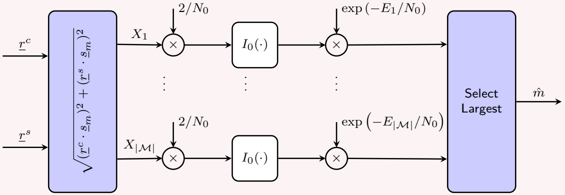


Module 12.1

Presentation Outline

Equal Energy Signals: Optimum Reception

General Structure



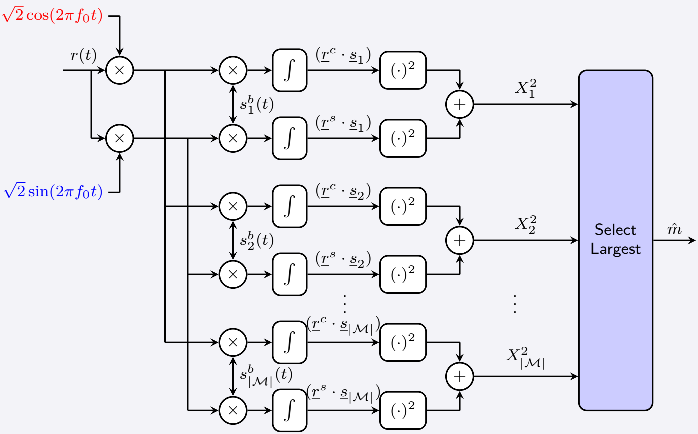
Equal Energy Signals

$$\hat{m}^{\mathsf{MAP}}(\underline{r}) = \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{ I_0\left(\frac{2\sqrt{(\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2}}{N_0} \right) \exp\left(-\frac{E_m}{N_0}\right) \right\}$$

$$= \underset{m\in\mathcal{M}}{\operatorname{argmax}} \left\{ (\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2 \right\}$$

Equal Energy Signals: Receiver Implementation

Correlation/Direct Receiver for Random Phase Equal-energy Signals



Summary Module 12.1

Take Home Messages

- Coherent vs. incoherent reception
- General structure of incoherent receiver: no knowledge of θ
- Equal energy signals are simpler to detect

Communication Theory (5ETB0)
Module 12.2

Alex Alvarado
a.alvarado@tue.nl

Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering
Eindhoven University of Technology, The Netherlands

www.tue.nl/ictlab/

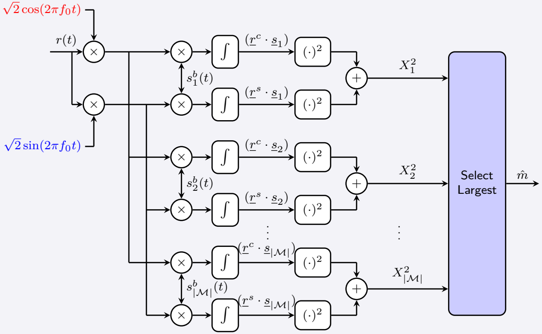
Module 12.2

Presentation Outline

Envelope Detection: Receiver Structure

Correlation/Direct Receiver for Random Phase Equal-energy Signals

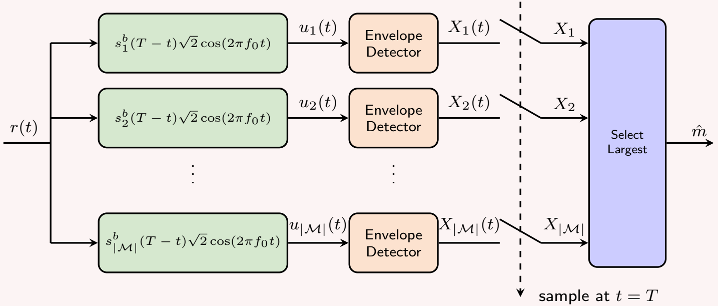
$$\hat{m}^{\mathrm{MAP}}(\underline{r}) = \underset{m \in \mathcal{M}}{\operatorname{argmax}} \left\{ X_m^2 \right\}, \qquad X_m \triangleq \sqrt{(\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2}$$



Envelope Detection: Receiver Structure

Envelope-detector receiver for equal energy signals with random phase

Received signal: $r(t) = s_m^b(t)\sqrt{2}\cos(2\pi f_0 t - \theta) + n_w(t)$



Questions

Is this an optimum receiver? Are the X_m here the same as before?

Envelope Detection: Optimally Proof (1/3)

Proof Sketch (Details in Sec. 12.4)

- Output of filters can be expressed as

$$u_m(t) = u_m^c(t)\underbrace{\cos(2\pi f_0 t)}_{\text{high-freq.}} + u_m^s(t)\underbrace{\sin(2\pi f_0 t)}_{\text{high-freq.}}$$

with

$$\underbrace{u_m^c(t)}_{\text{baseband}} \triangleq \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\cos(2\pi f_0 \alpha)\underbrace{s_m^b(T-t+\alpha)}_{\text{baseband}}\,d\alpha$$

$$\underbrace{u_m^s(t)}_{\text{baseband}} \triangleq \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\sin(2\pi f_0 \alpha)\underbrace{s_m^b(T-t+\alpha)}_{\text{baseband}}\,d\alpha$$

- Envelope detectors: (i) squaring signal, (ii) low-pass filter, and (iii) square root. The outputs are

$$X_m(t) = \sqrt{(u_m^c(t))^2 + (u_m^s(t))^2}$$

Envelope Detection: Optimally Proof (2/3)

Proof Sketch (Details in Sec. 12.4)

Note that

$$\begin{aligned}
(\underline{r}^c \cdot \underline{s}_m) = \sum_{i=1}^{N} r_i^c s_{mi} &= \sum_{i=1}^{N} \left(\int_{-\infty}^{\infty} r(t)\varphi_i(t)\sqrt{2}\cos(2\pi f_0 t)dt \right) s_{mi} \\
&= \int_{-\infty}^{\infty} r(t)\sqrt{2}\cos(2\pi f_0 t) \sum_{i=1}^{N} s_{mi}\varphi_i(t)dt \\
&= \int_{-\infty}^{\infty} r(t)\sqrt{2}\cos(2\pi f_0 t) s_m^b(t)dt.
\end{aligned}$$

Similarly,

$$(\underline{r}^s \cdot \underline{s}_m) = \int_{-\infty}^{\infty} r(t)\sqrt{2}\sin(2\pi f_0 t) s_m^b(t)dt.$$

Envelope Detection: Optimally Proof (3/3)

Proof Sketch (Details in Sec. 12.4)

$$
\begin{aligned}
u_m^c(t) &= \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\cos(2\pi f_0 \alpha) s_m^b(T-t+\alpha)d\alpha \Big|_{t=T} \\
u_m^c(T) &= \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\cos(2\pi f_0 \alpha) s_m^b(\alpha)d\alpha \\
u_m^s(t) &= \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\sin(2\pi f_0 \alpha) s_m^b(T-t+\alpha)d\alpha \Big|_{t=T} \\
u_m^s(T) &= \int_{-\infty}^{\infty} r(\alpha)\sqrt{2}\sin(2\pi f_0 \alpha) s_m^b(\alpha)d\alpha
\end{aligned}
$$

Thus, at $t = T$

$$
\begin{aligned}
u_m^c(T) &= (\underline{r}^c \cdot \underline{s}_m), \\
u_m^s(T) &= (\underline{r}^s \cdot \underline{s}_m),
\end{aligned}
$$

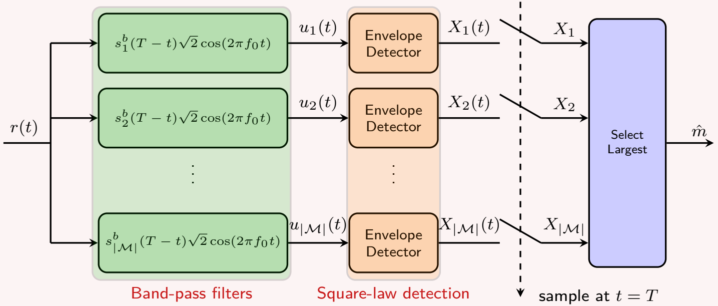
and therefore

$$X_m(T) = \sqrt{(\underline{r}^c \cdot \underline{s}_m)^2 + (\underline{r}^s \cdot \underline{s}_m)^2}.$$

Envelope Detection: Receiver Structure

Envelope-detector receiver for equal energy signals with random phase

Received signal: $r(t) = s_m^b(t)\sqrt{2}\cos(2\pi f_0 t - \theta) + n_w(t)$

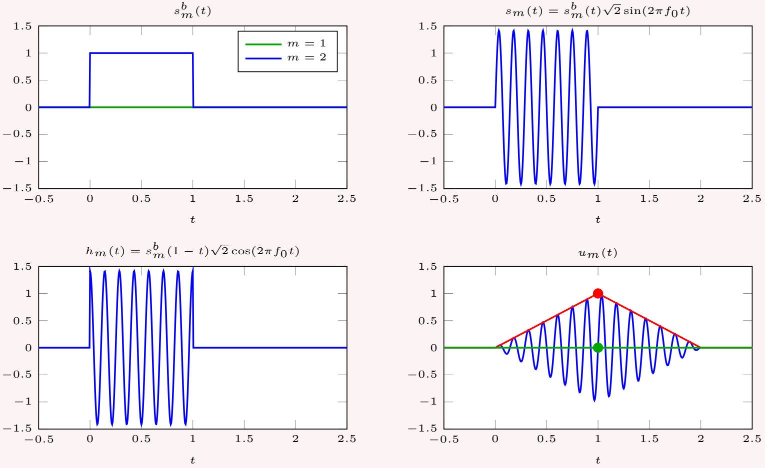


Questions

Is this an optimum receiver? Yes Are the X_m here the same as before? and Yes!

Envelope Detection: Example

Example 12.1 ($\theta = \pi/2$, $T = 1$)



Module 12.2

Presentation Outline

Part I Envelope Detection

Part II Error Probability

Error Probability for Two Orthogonal Signals (1/3)

Model and Assumptions

Signal vectors are $\underline{s}_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$, where

$$s_m^b(t) = \sum_{i=1}^{N} s_{mi}\varphi_i(t).$$

Two messages and two dimensions ($|\mathcal{M}| = N = 2$), equally likely. Waveforms are

$$\begin{aligned} s_1^b(t) &= \sqrt{E_s}\varphi_1(t), \text{ hence } \underline{s}_1 = (\sqrt{E_s}, 0), \\ s_2^b(t) &= \sqrt{E_s}\varphi_2(t), \text{ hence } \underline{s}_2 = (0, \sqrt{E_s}). \end{aligned}$$

All vectors are two-dimensional: $\underline{r}^c = \underline{s}_m \cos(\theta) + \underline{n}^c$, $\underline{r}^s = \underline{s}_m \sin(\theta) + \underline{n}^s$. Optimum receiver rule chooses $\hat{m} = 1$ if

$$(\underline{r}^c \cdot \underline{s}_1)^2 + (\underline{r}^s \cdot \underline{s}_1)^2 > (\underline{r}^c \cdot \underline{s}_2)^2 + (\underline{r}^s \cdot \underline{s}_2)^2$$
$$(r_1^c)^2 + (r_1^s)^2 > (r_2^c)^2 + (r_2^s)^2$$

Error Probability for Two Orthogonal Signals (2/3)

Error Probability Result (Sec. 12.5)

The error probability for an incoherent receiver for two equally likely orthogonal signals, both having energy E_s, is

$$P_{\mathrm{e}}^{in} = \frac{1}{2} \exp\left(-\frac{E_s}{2N_0}\right).$$

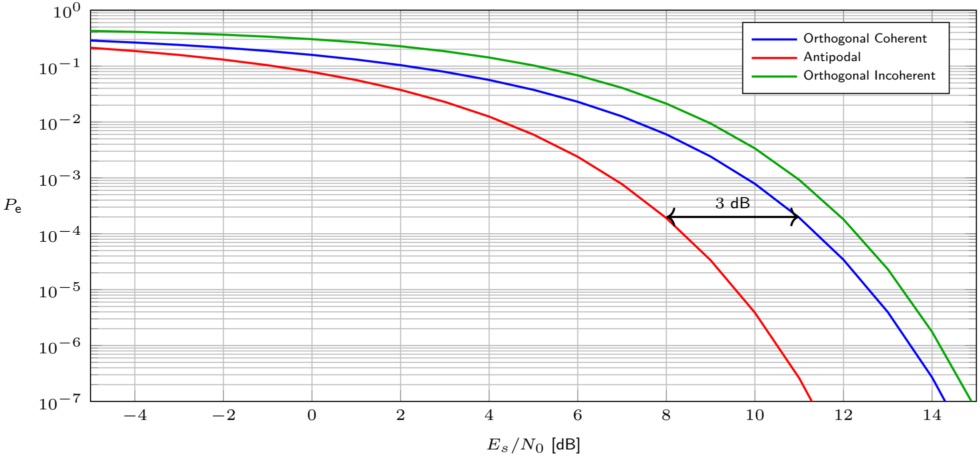
Comparison vs. Coherent and Antipodal

For coherent **(orthogonal)** reception and antipodal signaling we have obtained:

$$P_{\mathrm{e}}^{orth.} = Q\left(\sqrt{\frac{E_s}{N_0}}\right) \leq \frac{1}{2} \exp\left(-\frac{E_s}{2N_0}\right)$$

$$P_{\mathrm{e}}^{antip.} = Q\left(\sqrt{\frac{2E_s}{N_0}}\right)$$

Error Probability for Two Orthogonal Signals (3/3)



Pros and Cons of Incoherent Transmission

Pros and Cons

Advantages

- Simpler and cheaper implementation
- No need to track and compensate phase

Disadvantages

- Worse error probability for binary transmission and any SNR
- More than 3 dB loss compared to antipodal signaling

Summary Module 12.2

Take Home Messages

- Optimum incoherent receiver can be implemented easily
- Analysis based on building-block waveforms
- Error probability analysis: performance degradation